UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

Metodologias para análise de experimentos de Dual RNA-seq: aplicação da melhor estratégia no estudo das interações entre Azospirillum brasilense e milho durante a inibição da produção de Ácido Indol-3-Acético pela planta

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau Doutor em Ciências (Genética e Biologia Molecular).

Eliandro Espindula

Orientadora: Profa. Dra. Luciane M. P. Passaglia

Porto Alegre, agosto de 2019

Este trabalho foi desenvolvido nas Universidades Federais do Rio Grande do Sul (UFRGS) e do Paraná (UFPR), com financiamento da CAPES, CNPq/INCT-FBN e Fundo Newton (Brasil/Reino Unido).

Dedico este trabalho a Deus, que sempre guiou e iluminou meus passos, ao me mostrar sempre quais os melhores caminhos a seguir. Ele me segurou em seus braços quando fraquejei no caminho e sempre, quando eu mais precisei de apoio, veio em meu socorro, através de anjos maravilhosos que a gente por aqui chama de amigos.

Agradecimentos

Primeiramente, a Deus, por me manter sempre no caminho certo.

À Universidade Federal do Rio Grande do Sul, por me proporcionar a estrutura física e logística que me permitiram desenvolver este trabalho.

À minha orientadora, Luciane M. P. Passaglia, que além de minha orientadora, foi também minha conselheira, me trazendo de volta para o caminho quando eu parecia perder o foco. Me mostrando qual melhor direção quando tantas se fizeram presentes no curso dessa pesquisa. Me trazendo paz quando as coisas pareciam não dar certo, me lembrando que em ciência é assim, é tentativa e erro, até a gente conseguir fazer certo para conseguir produzir o ambiente científico certo para a investigação que pretendemos.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular, por proporcionar a oportunidade de poder participar do curso de doutorado oferecido por este órgão e, assim, me conceder a chance de ter acesso a todo o conhecimento disponível neste por meio dos renomados professores que fazem parte deste programa.

À CAPES, ao INCT/FBN e ao Fundo Newton por proporcionarem os recursos necessários para a execução deste trabalho.

Aos meus pais, pelo Amor a mim dedicado e pela lição de vida de que somente com trabalho, dedicação e fé nós podemos alcançar nossos objetivos na vida.

Aos meus pais, Alberi Espindula e Leda Maria Espindula, meu irmão, Heitor Espindula, à tia Noélcia Espindula Bressan, ao tio Valter Bressan, à Tia Nera Espindula Cerva, ao tio Manoel (Maneco) Cerva, à tia Nelza (Dinda) Espindula Schieck e aos primos, frutos dessas uniões, por todo Amor e por me ensinarem o que significa ser uma família.

Aos amigos que sempre me apoiaram nos bons e maus momentos, me aconselhando e mostrando os caminhos em noites escuras e mal iluminadas.

À amiga, hoje ausente, Edilena Reis Sperb, que não foi somente uma parceira de trabalho, foi muito mais. Foi amiga daquelas que a gente sente falta, foi daquelas que olha no teu olho e te diz o que tu precisas ouvir, não o que você quer ouvir. Foi amiga parceira nas “indiadas”, como a gente diz aqui no RS. Foi a parceira daqueles dias que a gente somente quer ficar sentado, olhando para o longe sem dizer uma palavra, eu na minha cerveja e você na tua eterna Gin Tônica. Foi a parceira de trabalho, pois metade do que ele é hoje eu devo à tua visão de conseguir ver nele a potencialidade que eu não via.

À amiga Jacque(line) Barcelos pela incomparável e impagável paciência, amor e carinho. Tu és daquelas amigas que a gente conta nos dedos, que leva para a vida. Tu foste e ainda és um farol para mim, pois és dona de uma sabedoria e de uma visão de vida incríveis, que, apesar das dificuldades da vida, sempre consegue ver o melhor dela. Tu onde passas é luz, assim como tu és na minha. Te sou eternamente grato por tudo!

Ao amigo Rafael Rauber por todo o carinho, parceria em todos os momentos. Já defendestes e já estás a ensinar e, ainda assim, sempre encontras tempo para o amigo aqui. Seja por conselhos, seja por risadas ou aquela parceria na cerveja (onde eu tinha que convencer todo mundo que não sabia quem era tua namorada, porque vocês não queriam). Convivemos por mais ou menos um ano. Mas os laços que nos unem são aqueles de duas pessoas que parecem se conhecer de outra vida dado tamanho da cumplicidade na amizade.

Aos meus eternos mestres nas ciências e na vida. Até eu chegar aqui eles eram os prof[s] Umberto Euzebio (UnB), Silviene Oliveira (UnB), Zulmira M. Lacava (UnB), Loreny Giugliano (UnB), e a pesquisadora da Embrapa Rose Gomes Monnerat Sólon de Pontes (CENARGEN) e, hoje, essa lista aumentou! Durante esses quatro anos que estive aqui, tive a honra e o prazer de conviver com as professoras Márcia Pinheiro Margis, Eliane Kaltchuk e a minha orientadora Luciane M.P. Passaglia. Assim, obrigado amados mestres que, com seu conhecimento e amizade, me ensinaram o que é ser um cientista: Pessoa que consegue, ao se deparar com o desconhecido, analisar a situação a ele posta e, por meio da experimentação e dos conhecimentos disponíveis, alcançar a explicação para a situação. Levarei sempre comigo todas as lições aprendidas até hoje e aquelas que vocês, todos os dias, continuam a me ensinar.

À vida, que me ensinou que o melhor é sempre seguir em frente, sem deixar que as pedras e buracos da estrada atrapalhem nossa caminhada e que, acima de tudo, devemos aprender com nossos acertos e, principalmente, com nossos erros.

“Comece fazendo o que é necessário, depois o que é possível e, de repente, você estará fazendo o impossível.”

(São Francisco de Assis)

Índice

Abreviaturas:

AIA/IAA = Ácido Indol – 3 –Acético / Indole – 3 – Acetic Acid

AIP / IPA = Ácido Indol – 3 – Piruvato / Indole – 3 – Pyruvate

ABA = *Abscysic Acid* (Ácido Abscísico)

DEGs = *Differentially Expressed Genes* (Genes Diferencialmente Expressos)

DNA = *Deoxyribonucleic Acid* (Ácido Desoxirribonucleico)

GA = Giberelina

GO = *Gene Ontology* (Ontologia Gênica)

PGPB = *Plant-Growth Promoting Bacteria* (Bactérias Promotoras de Crescimento Vegetal)

RNA = *Ribonucleic Acid* (Ácido Ribonucleico):

mRNA= RNA mensageiro

rRNA = RNA ribossômico

tRNA = RNA transportador

TF = *Transcription Factors* (fatores de transcrição)

Trp = Triptofano

Resumo

Uma forma de elevar a produção vegetal sem aumentar a área plantada é através do uso de bactérias promotoras de crescimento vegetal. Entre essas bactérias estão as do gênero *Azospirillum*, encontradas na rizosfera de gramíneas e cereais, tanto em climas tropicais como em climas temperados. As principais características das bactérias desse gênero são a capacidade de fixar nitrogênio e a de produzir fitormônios, estimulando, assim, o crescimento das plantas. O presente estudo teve como objetivo melhorar a compreensão sobre a interação de *Azospirillum* e milho, em nível genético, em relação à ação do ácido indol-3-acético (AIA) nas plantas. Esse fitormônio é muito importante para o crescimento vegetal, principalmente através da estimulação da produção de raízes secundárias. Para esta finalidade, foi aplicado nas plantas o composto químico Yucasin [5-(4-chlorofenil) -4H-1,2,4-triazol-3-tiol], que é um potente inibidor de uma das enzimas da principal via de produção de AIA pelas plantas, a YUCCA. Plântulas de milho, inoculadas ou não com *Azospirillum brasilense* FP2, foram submetidas ao tratamento com Yucasin por 5 horas. Amostras de raízes de cada condição experimental foram utilizadas para a extração de RNA e para a construção das respectivas bibliotecas de cDNA. Uma vez que a metodologia utilizada foi de *Dual RNA-seq*, fez-se necessário o aprimoramento da forma de análise dos dados gerados após o sequenciamento das bibliotecas de cDNA, constituídas de amostras de RNA de ambos os organismos. Apesar de a maioria dos trabalhos de *Dual RNA-seq* utilizar a análise sequencial dos dados, a análise combinada, na qual se compara simultaneamente os dados, parece ser a mais indicada. Nesse tipo de análise, as bibliotecas são alinhadas contra um arquivo composto pelos genomas de ambos os organismos de forma concatenada, antes da etapa de mapeamento das sequências contra os respectivos genomas anotados. Para verificar a eficácia da metodologia combinada em relação à sequencial foram utilizadas bibliotecas de cDNAs disponíveis em bancos públicos e bibliotecas oriundas de um experimento prévio realizado em nosso laboratório. A metodologia sequencial atribuiu um número maior de sequências ao primeiro genoma utilizado (devido ao mapeamento cruzado) do que a metodologia combinada. Também, a quantidade de mapeamento cruzado foi menor na análise combinada do que na sequencial, evidenciando a importância da utilização da análise combinada em experimentos de *Dual RNA-seq*, a fim de se evitar a perda de informações, principalmente em relação ao organismo procariótico. A análise combinada foi aplicada para analisar as bibliotecas de cDNA do experimento *Azospirillum*-milho, na presença ou não de Yucasin. Com essa metodologia foi possível observar em milho genes diferencialmente expressos que codificam proteínas da via de resposta ao ácido abscísico (ABA) e fatores de transcrição evolvidos nas respostas a estresses bióticos e abióticos. Esses resultados indicam que a planta respondeu à presença da bactéria e do Yucasin. Dados fisiológicos mostraram que o AIA produzido pela bactéria foi suficiente para recuperar o fenótipo das plantas submetidas ao tratamento com Yucasin. Com relação à bactéria, foram identificados genes diferencialmente expressos que codificam proteínas que participam do transporte transmembrana e da biossíntese de terpenoides.

Abstract

One way to increase crop yields without increasing planted area is through the use of plant growth-promoting bacteria. These include bacteria of the genus *Azospirillum* found in the grass and cereal rhizosphere in both tropical and temperate climates. The main characteristics of bacteria of this genus are the ability to fix nitrogen and to produce phytohormones, thus stimulating plant growth. The present study aimed to improve the genetic understanding of the interaction of *Azospirillum* and maize concerning the action of indole-3-acetic acid (IAA) on plants. This phytohormone is very important for plant growth, primarily through the stimulation of secondary root production. For this purpose, the chemical compound Yucasin [5-(4-chlorophenyl) -4H-1,2,4-triazol-3-thiol], which is a potent inhibitor of one of the enzymes (YUCCA) of the major plant production AIA pathway, was applied to the plants. Corn seedlings, inoculated or not with *Azospirillum brasilense* FP2, were submitted to treatment with Yucasin for 5 hours. Root samples from each experimental condition were used for RNA extraction and construction of the respective cDNA libraries. Since the methodology used was Dual RNA-seq, it was necessary to improve the analysis of the data generated after the sequencing of cDNA libraries, consisting of RNA samples from both organisms. Although most Dual RNA-seq works use sequential data analysis, combined analysis, in which data are simultaneously compared, seems to be the most appropriate. In this type of analysis, libraries are aligned against a file composed of the genomes of both organisms in a concatenated manner prior to the sequence mapping step against their annotated genomes. To verify the effectiveness of the combined methodology in relation to the sequential, we used libraries of cDNAs available in public banks and libraries from a previous experiment performed in our laboratory. The sequential methodology assigned more sequences to the first genome used (due to cross-mapping) than the combined methodology. Also, the amount of cross-mapping was lower in the combined than in the sequential analysis, evidencing the importance of using the combined analysis in Dual RNA-seq experiments, in order to avoid information loss, especially about the prokaryotic organism. The combined analysis was applied to analyze the cDNA libraries of the *Azospirillum*-maize experiment, with or without Yucasin. With this methodology, it was possible to observe in maize differentially expressed genes that encode abscisic acid (ABA) response pathway proteins and transcription factors involved in responses to biotic and abiotic stresses. These results indicate that the plant responded to the presence of bacteria and Yucasin. Physiological data showed that the AIA produced by the bacterium was sufficient to recover the phenotype of plants subjected to treatment with Yucasin. For bacteria, differentially expressed genes that encode proteins that participate in transmembrane transport and terpenoid biosynthesis were identified.

1. Revisão bibliográfica

1.1. O Milho (*Zea mays*)

O milho (*Zea mays*) é uma planta originada no México e que possui variedades adaptadas a diversos ambientes terrestres, sendo uma das culturas de maior importância mundial (Fornasieri Filho, 1992). O Brasil, segundo dados da Organização das Nações Unidas para a Alimentação e Agricultura (FAO, 2016), apresentou, em 2016, a terceira maior produção mundial de milho, atrás somente de Estados Unidos e China. O milho cultivado no Brasil é insumo para diversos produtos e seu principal uso (entre 70 e 80%) é como ração para suínos e aves (Duarte *et al.* 2015) . Tal fato é importante na medida em que contribui, indiretamente, com a alimentação humana, pois, ao ser usado como alimento para suínos e aves, esse milho ingerido por estes animais auxilia na formação das reservas proteicas dos mesmos que serão, após o abate, consumidas por humanos. O milho, além de contribuir de forma indireta na alimentação humana, contribui também de forma direta ao ser consumido *in natura* ou processado.

A cultura do milho apresenta demandas nutricionais específicas para seu sucesso e para que apresente uma alta produtividade são necessários em média 104 Kg de nitrogênio por hectare (Hungria *et al.* 2010). O nitrogênio é um elemento essencial na constituição de todas as proteínas celulares. Dessa forma, a sua disponibilidade no ambiente em uma forma que possa ser prontamente absorvida e utilizada pelos seres vivos é primordial à manutenção da vida. Assim, o nitrogênio, juntamente com a água, é tido como um fator limitante de crescimento, tanto de plantas em ambientes naturais, quanto em plantas cultivadas, como o milho (*Zea mays*). Entretanto, a maior parte dos estoques de nitrogênio do solo se encontra em uma forma não disponível às plantas (inorgânica), dependendo de micro-organismos presentes no solo para a conversão deste nitrogênio nas suas formas minerais orgânicas (NH_4^+ e NO_3^-). Estas formas minerais são as formas prontamente absorvidas pelas plantas (Poletto, 2004). No entanto, a mineralização do nitrogênio não é suficiente para sustentar as altas demandas deste nutriente pelas culturas comerciais, sendo necessário realizar a adição de nitrogênio mineral (Weber and Mielniczuk, 2009), o qual, quando adicionado ao solo, não é imediatamente absorvido pelas plantas (Machado *et al.* 1998). Em condições normais, as plantas absorvem menos de 50% do nitrogênio mineral adicionado na forma de fertilizantes. Isso ocorre porque parte do nitrogênio é perdida por lixiviação deste na forma de nitrato, volatilização da amônia e pela emissão de N_2 e óxidos de nitrogênio (Bredemeier and Mundstock, 2000). O nitrogênio que se perde por lixiviação acaba por se tornar uma das principais fontes de poluição dos sistemas vegetais. Além da poluição causada pelo nitrogênio que é lixiviado, há também a poluição gerada pela produção industrial dos fertilizantes, a qual envolve

o consumo de combustíveis fósseis e a emissão de gases para a atmosfera (Weber and Mielniczuk, 2009).

1.2.Bactérias Promotoras de Crescimento Vegetal (PGPB):

Bactérias promotoras de crescimento vegetal (PGPB – *Plant-Growth Promoting Bacteria*, em inglês) constituem um grupo de micro-organismos benéficos às plantas devido à sua capacidade de colonizar a superfície de raízes, a rizosfera, a filosfera e os tecidos internos da planta (Hungria *et al.* 2010) e de promover o crescimento vegetal (Verna *et al.* 2010). Essas bactérias são capazes de estimular um maior crescimento vegetal por meio da fixação biológica de nitrogênio (FBN), da produção de fitormônios, principalmente ácido indol-3-acético (AIA), vitaminas e fatores de crescimento que estimulam o crescimento e a produção das plantas. Acredita-se que estas bactérias estimulam o crescimento vegetal por meio da combinação de vários desses mecanismos (Babalola, 2010; Hungria *et al.* 2010; Bashan and de-Bashan, 2010).

Dentre os fitormônios, o mecanismo de produção de AIA pelas PGPB tem sido um dos mais estudados (Yue *et al.* 2014; Baudoin, 2010; Spaepen *et al.* 2007; Spaepen and Vanderleyden, 2010). A principal via de produção desse hormônio em bactérias, assim como em plantas, usa o aminoácido triptofano (Trp) como precursor da síntese de AIA (Spaepen *et al.*2007) (Figura 1). Apesar dessa similaridade, não são conhecidos homólogos em PGPB das enzimas envolvidas na principal via de produção em plantas. Tal fato indica que a via de produção de AIA evoluiu de forma independente em bactérias e plantas terrestres (Yue *et al.* 2014). Essa evolução convergente aponta a importância dessa substância em ambos os grupos. O AIA é usado pelas bactérias como uma molécula sinalizadora que permite às bactérias estimular o crescimento radicular e exsudação de carboidratos pelas plantas e, assim, criar um ambiente favorável para seu crescimento. Além disso, essa molécula está envolvida no processo de *quorum sensing* das bactérias. *Quorum sensing* é o uso de moléculas sinalizadoras capazes de se difundir pela membrana na comunicação entre as células (Crépin et al., 2012). Isso permite que as bactérias percebam a densidade da população e controlem suas atividades com base nela. (Crépin et al., 2012; Yue *et al.* 2014; Duca *et al.* 2014; Spaepen *et al.* 2007).

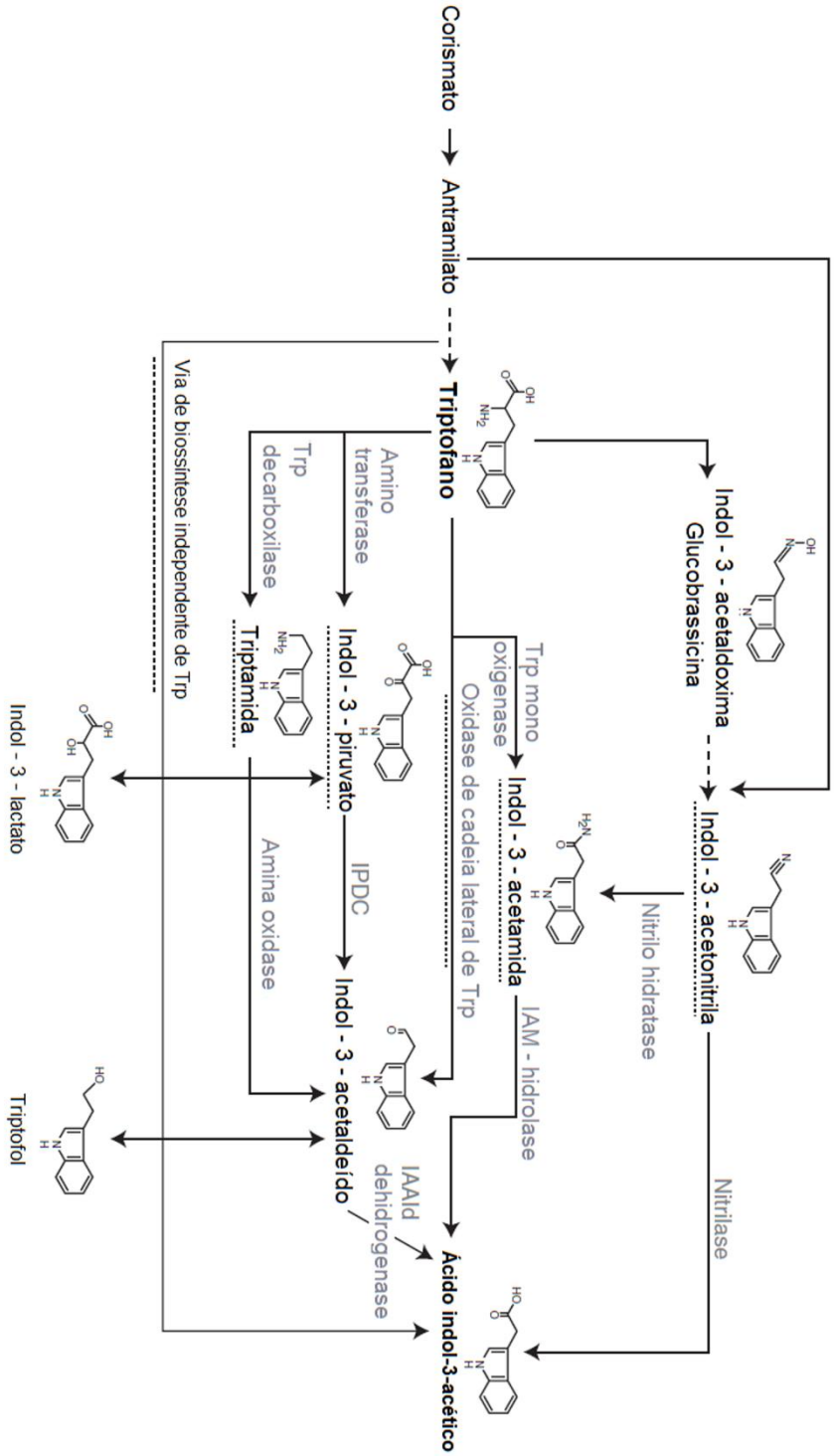


Figura 1. Visão geral das vias de produção de auxinas pelas bactérias. Os nomes sublinhados com linhas pontilhadas se referem aos nomes das vias de síntese. Os nomes em cinza se referem aos nomes das enzimas envolvidas nas reações. IAAId, Indol–3–acetaldeído; IAM, Indol–3–acetamida; IPDC, indol–3–piruvato descarboxilase; Trp, Triptofano. Adaptada de Spaepen and Vanderleyden (2010).

Por fim, as raízes das plantas excretam substâncias capazes de interagir com o sistema de *quorum-sensing* das bactérias, atraindo-as para a rizosfera. As PGPB, ao chegarem à rizosfera, colonizam esta região e começam a produzir substâncias que favorecem o crescimento das plantas (fitormônios, vitaminas, aminoácidos, fatores de crescimento, regulação da produção de etileno, entre outras) (Babalola, 2010).

Os seguintes gêneros bacterianos possuem membros que fazem parte do grupo das PGPB: *Acetobacter, Acinetobacter, Alcaligenes, Arthrobacter, Azoarcus, Azospirillum, Azotobacter, Bacillus, Beijerinckia, Burkholderia, Derxia, Enterobacter, Gluconacetobacter, Herbaspirillum, Klebsiella, Ochrobactrum, Pantoae, Pseudomonas, Rhodococcus, Serratia, Stenotrophomonas* e *Zoogloea (*Verna *et al. 2010)*. Com exceção das bactérias do gênero *Rhizobium*, as mais estudadas dentre as PGPB conhecidas são as pertencentes ao gênero *Azospirillum* (Bashan and de-Bashan, 2010)

1.3.Azospirillum spp.:

Azospirillum é um gênero de bactéria cujos representantes são encontrados em praticamente todos os locais do planeta (Hungria *et al.* 2010). São bactérias gram-negativas de vida livre que podem ser isoladas da rizosfera de gramíneas e de cereais, cultivados tanto em climas tropicais como em climas temperados (Steenhoudt and Vanderleyden, 2000). Essas bactérias possuem como características principais serem capazes de fixar o nitrogênio gasoso (N_2) e produzirem fitormônios (Cassán *et al.* 2009). *Azospirillum* spp. são capazes de fixar o nitrogênio gasoso (N_2) em condições microaeróbicas e o converter em amônia (NH_4^+) (Steenhoudt and Vanderleyden, 2000), que é uma das formas minerais do nitrogênio que é prontamente absorvida pelas plantas (Fornasieri, 1992). Tal forma mineral é excretada por estas bactérias, o que torna o nitrogênio disponível às plantas.

A fixação biológica de nitrogênio ocorre por meio do complexo da nitrogenase. Esse complexo é formado por duas metaloproteínas: a dinitrogenase redutase (ou Fe-proteína, produto do gene *nifH*) e a dinitrogenase (Mo-Fe proteína, que contém um cofator ferro-molibdênio, FeMoco, produto dos genes *nifDK*) (Steenhoudt and Vanderleyden, 2000). O complexo da nitrogenase reduz o N_2 à NH_4^+, como mostra a equação abaixo (Huergo, 2006):

$$\mathbf{N_2 + 10H^+ + 8e^- + 16Mg.ATP \rightarrow 2}\boldsymbol{NH_4^+ + H_2 + 16Mg.ADP + 16Pi}$$

A Fe-proteína se liga a duas moléculas de MgATP, o que altera a conformação desta, e, assim, permite que a Fe-proteína se ligue à Mo-Fe proteína. Uma vez formado o complexo da nitrogenase, a Fe-proteína transfere um elétron para o cofator Fe-Mo da Mo-Fe proteína. Durante

a transferência do elétron ocorre a hidrólise das duas moléculas de MgATP, liberando dois fosfatos inorgânicos. O MgADP permanece ligado à Fe-proteína oxidada e o complexo da nitrogenase é desfeito. A Fe-proteína é, então, reduzida por um doador de elétrons e os dois MgADP são substituídos por outras duas moléculas de MgATP, reiniciando o processo. Esta sequência de reações é repetida até que sejam transferidos elétrons suficientes para que o N_2 possa ser reduzido à NH_4^+ (Huergo, 2006), em número de oito (Steenhoudt and Vanderleyden, 2000, Huergo, 2006).

Após a redução do N_2 a NH_4^+ ou da obtenção de amônia a partir do meio, a bactéria assimila essa amônia por meio da produção de glutamina e glutamato, compostos que irão atuar como doadores de nitrogênio nas reações celulares (Arcondeguy *et al.* 2001).

Segundo Machado *et al.* (1998), em experimentos em que se inoculou plantas de milho com uma mistura de bactérias promotoras de crescimento vegetal (*A. amazonense*, *A. lipoferum* e *Herbaspirillum seropedicae*) em dois regimes de adubação nitrogenada (10 kg ha^{-1} e 100 kg ha^{-1}), em que se consideraram as médias dos tratamentos com e sem inoculação, houve aumento no número de grãos, no conteúdo de nitrogênio nos grãos e no total de nitrogênio presente nas plantas. Conforme resultados obtidos por Junior *et al.* (2008), a inoculação com a bactéria *Azospirillum amazonense* em dois híbridos de milho, sob três regimes de adubação nitrogenada (100% NH_4^+, 100% NO_3^- e sem Nitrogênio), produziu um aumento na quantidade de matéria seca e nitrogênio nas raízes das plantas inoculadas, quando comparadas com as não inoculadas.

Além de fixar o nitrogênio atmosférico, bactérias do gênero *Azospirillum* também produzem fitormônios (ácido indol-3-acético – AIA; giberelinas – GA_3 – Ácido Abscísico, etileno e poliaminas) que se acredita serem responsáveis, juntamente com a fixação de nitrogênio, por estimular o aumento na massa seca final das plantas (Bashan and de-Bashan, 2010; Perrig *et al.* 2007). Por essa razão, desde a sua redescoberta na metade da década de 1970, pela pesquisadora Johanna Döbereiner e seus colaboradores (Bashan and de-Bashan, 2010), a interação entre a planta e as bactérias do gênero *Azospirillum* e sua capacidade de fixar o nitrogênio atmosférico tem sido amplamente estudada. Desde então, os trabalhos objetivaram elucidar o funcionamento das vias metabólicas desta bactéria de forma a entender os processos realizados por ela para fixar o nitrogênio atmosférico e para produzir fitormônios (Elmerich *et al.* 1997; Cassán *et al.* 2001; Santos *et al.* 2010; Molina-Favero *et al.* 2008).

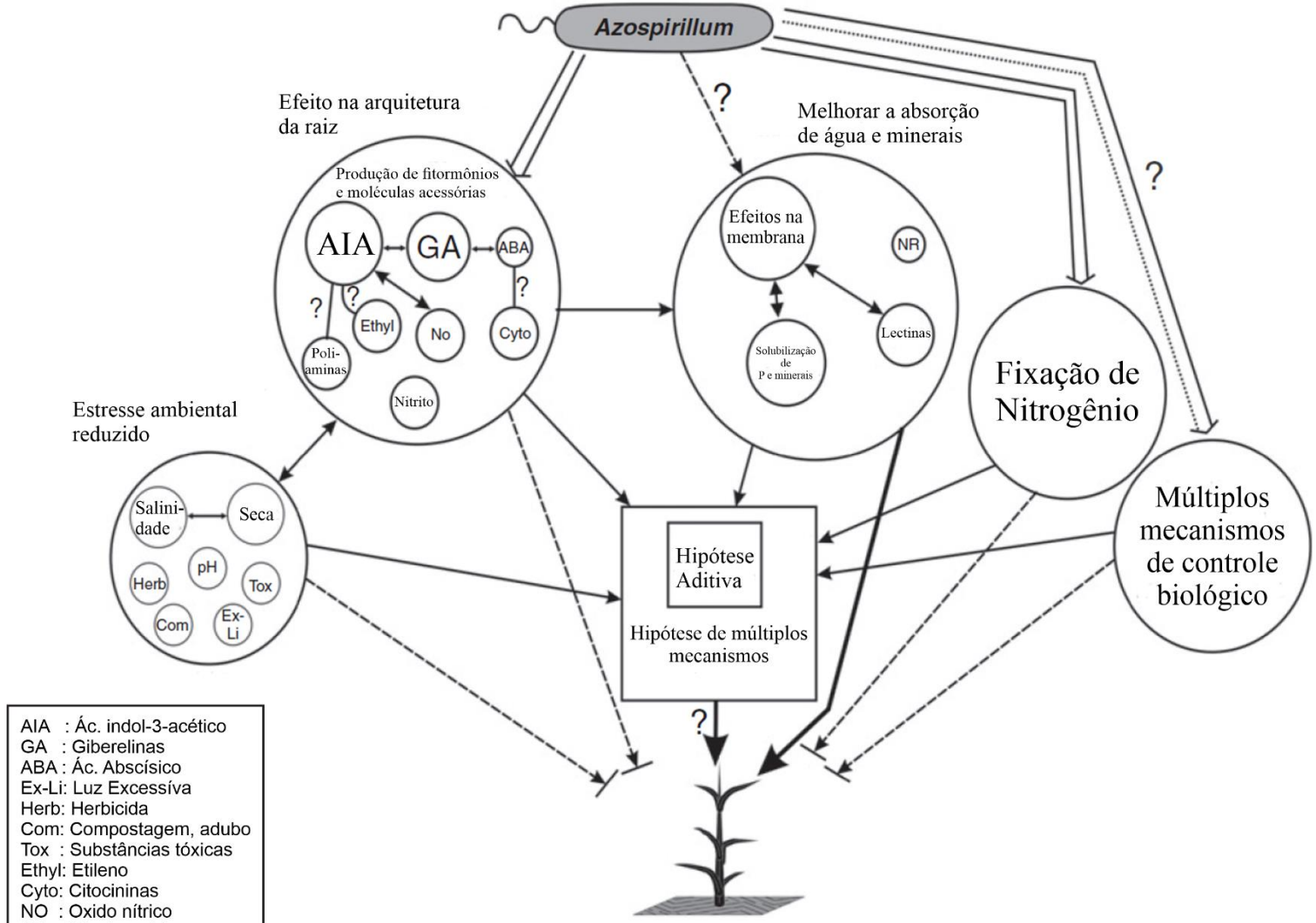

Azospirillum
Efeito na arquitetura da raiz
Melhorar a absorção de água e minerais
Produção de fitormônios e moléculas acessórias
AIA
GA
ABA
Ethyl
No
Cyto
Poli-aminas
Nitrito
Efeitos na membrana
NR
Lectinas
Solubilização de P e minerais
Fixação de Nitrogênio
Estresse ambiental reduzido
Salini-dade
Seca
pH
Herb
Tox
Com
Ex-Li
Hipótese Aditiva
Hipótese de múltiplos mecanismos
Múltiplos mecanismos de controle biológico
AIA : Ác. indol-3-acético
GA : Giberelinas
ABA : Ác. Abscísico
Ex-Li: Luz Excessíva
Herb: Herbicida
Com: Compostagem, adubo
Tox : Substâncias tóxicas
Ethyl: Etileno
Cyto: Citocininas
NO : Oxido nítrico
NR : Nitrato redutase

Figura 2. Mecanismos pelos quais *Azospirillum* spp. pode melhorar o crescimento das plantas e suas possíveis interações agrupadas como processos biológicos. Círculos representam processos contendo dados experimentais. Quadrados representam teorias. O tamanho do círculo representa a importância relativa de acordo com os dados atuais. Seta sólida: mecanismo(s) que pode(m) criar totalmente a promoção de crescimento observada; seta pontilhada: mecanismo(s) que pode(m) apenas explicar parcialmente a promoção de crescimento observada. Setas simples: interações comprovadas entre diferentes mecanismos; seta de linha dupla: produção direta de moléculas ou processos realizados pela célula bacteriana. "?" : não provado ainda, ou provado parcialmente. Adaptado de Bashan e de-Bashan, 2010.

É bem conhecido que o triptofano (Trp) é o precursor da síntese de AIA em *Azospirillum brasilense*, sendo a enzima indol–3–piruvato descarboxilase (IPDC) a principal responsável por esta síntese, e que baixas concentrações de Trp favorecem a síntese do AIA (Reynders e Vlassak, 1979; Zimmer *et al.* 1998; Malhotra e Srivastava, 2006; Malhotra e Srivastava, 2008). Segundo Ona *et al.* (2005), a biossíntese de AIA ocorre quando há limitação nutricional (baixas concentrações de nitrogênio e de fontes de carbono) e a bactéria se encontra associada com as raízes de plantas, o que resulta em baixos níveis de oxigênio e na presença de Trp. Segundo Ribaudo *et al.* (2006), o AIA produzido por *A. brasilense* durante a interação com o tomate, juntamente com o AIA produzido pela planta, é capaz de estimular a síntese de etileno pela planta em níveis que não inibem a elongação da raiz e a formação de raízes laterais. Por fim, segundo Duca *et al.* (2014), o AIA é uma molécula de sinalização usada na interação planta-bactéria que sustenta a relação simbiótica entre plantas e bactérias benéficas que evoluiu ao longo dos tempos.

Outros estudos apontaram a capacidade de *A. brasilense* de produzir as formas ativas de giberelinas (GA_1 e GA_3) (Bottini *et al.* 2004; Perrig *et al.* 2007; Salazar-Cerezo *et al.* 2018) e de realizar a 3-*β*-hidroxilação das formas inativas das 3-*β*-deoxi GA presente em raízes (Cassán *et al.* 2001)

1.4.Auxinas

Nos últimos anos, muitos pesquisadores têm estudado os mecanismos de ação dos fitormônios no crescimento e desenvolvimento vegetal. Entre os principais grupos de fitormônios estudados estão as auxinas (especialmente o ácido indol-3-acético, AIA) (Blakeslee *et al.* 2005; Carraro *Et al.* 2006; Chen and Gallie, 2010; Hirano *et al.* 2008; Simon and Petrasek, 2011; Burg and Thimann, 1959; Ljung *et al.* 2001; Ljung *et al.* 2005; Tanimoto *et al.* 1995). As Auxinas são um grupo de hormônios responsáveis por regularem vários aspectos do desenvolvimento vegetal, como o crescimento e a diferenciação celular, estabelecimento da dominância apical, diferenciação do xilema, supressão da abscisão na formação dos meristemas apical radicular e caulinar (Bishopp

et al. 2006; Yue *et al.* 2014). Segundo, Simon and Petrasek (2011), o mecanismo de ação da auxina inclui respostas rápidas que não envolvem a expressão gênica e respostas mais lentas que requerem a expressão de genes regulados pela auxina, mediada pelas proteínas F-box.

As auxinas são sintetizadas, principalmente, nos meristemas apicais e em folhas jovens (Tromas and Perrot-Rechenmann, 2010) e, também, no meristema apical das raízes, especialmente durante a fase de elongação da raiz. Entretanto, a raiz ainda depende da auxina produzida nas partes aéreas da planta (Ljung *et al.* 2001; Ljung *et al.* 2005). Nas plantas, a principal via de produção do AIA é a partir do aminoácido triptofano (Trp). Ao longo dos anos, várias rotas de produção de AIA a partir de Trp foram propostas (Zhao, 2010). Nesses estudos, as enzimas pertencentes às famílias das Triptofano-aminotransferases de *Arabidopsis* (TAA) e as das YUC flavino-mono-oxigenases eram posicionadas em rotas metabólicas diferentes. Entretanto, estudos recentes indicaram que essas duas enzimas fazem parte da principal via de produção de AIA nas plantas. A TAA converte o triptofano em Indol-3-piruvato (IPA), que é convertido em AIA pela YUC (Figura 3) (Zhao, 2012; Zhao, 2014; Won *et al.* 2011; Mashiguchi *et al.* 2011; Yue *et al.* 2014). Por fim, foi observado por Nishimura *et al.* (2014) que a substância conhecida como Yucasin [5-(4-clorofenil)-4H-1,2,4-triazol-3-tiol] é um inibidor competitivo de IPA, impedindo a descarboxilação deste pela YUC. Por inibir de forma sinérgica a produção de AIA por esta via e o crescimento da planta, o Yucasin pode ser usado em estudos de modulação da produção de AIA pela planta ao longo do tempo (Zhao, 2014).

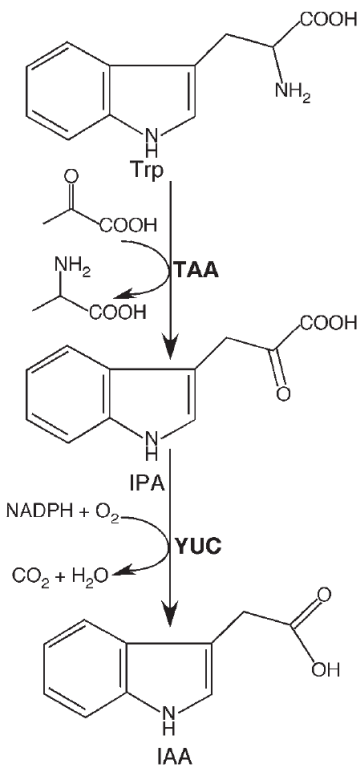


Figura 33. Via de biossíntese de auxinas dependente de triptofano em plantas. A primeira etapa é catalisada pelas TAAs que transferem o grupo amino do Trp para um ácido ceto-alfa, como o piruvato, para formar o IPA e outro aminoácido. A segunda etapa é uma reação dependente de NADPH e oxigênio que tem como produtos CO_2, H_2O e AIA. Essa reação é catalisada pelas YUC flavina-mono-oxigenases. Adaptada de Zhao (2012).

Além das plantas, bactérias e fungos também são capazes de produzir AIA a partir de vias dependentes de Trp (Zhao, 2010; Yue *et al.* 2014). Nesse contexto, não se sabe se e como as vias metabólicas de plantas e bactérias estão relacionadas, pois foram achadas tanto similaridades quanto diferenças em nível molecular nas vias de produção de AIA por plantas e bactérias. Em relação à biossíntese desse hormônio por fungos, essa parte continua pendente de mais estudos. Assim, apesar da importância desse hormônio no desenvolvimento das plantas, nosso conhecimento sobre a evolução das vias biossíntese de auxinas pelas plantas continua limitada (Yue *et al. 2014*).

Muito desse desconhecimento sobre a evolução da biossíntese das auxinas foi causada pela falta de conhecimento da via de produção de AIA por si só. Muitos caminhos foram sugeridos ao longo dos anos (Yue *et al.* 2014). Entretanto, a recente descoberta da via TAA/YUC para produção de AIA a partir de Trp e que essa é a principal via de produção em plantas (Won *et al.* 2011; Mashiguchi *et al.* 2011) permitiu uma nova análise sobre a origem da via de produção das auxinas nas plantas. Um fato interessante é que em algas marrons foram encontrados vários genes homólogos aos genes de planta para a biossíntese de AIA. Isso sugere a possibilidade desse hormônio ser produzido em várias linhagens de algas (Yue *et al.* 2014; ver Refs 15-16 dessa revisão para maiores detalhes), e inclusive já era conhecido que homólogos dos genes envolvidos no transporte das auxinas e o do seu receptor estavam presentes nas algas verdes (Yue, 2017). Entretanto, outros componentes da maquinaria de resposta às auxinas, como o Fator de Resposta às Auxinas (ARF), são específicos das plantas terrestres. Apesar disso, vários genes homólogos da via biossintética de AIA dependente de Trp foram identificados nas algas verdes. Tal fato indica

uma origem mais antiga da via de produção de AIA. Assim, tomando por base a via metabólica composta pelas famílias gênicas das Triptofano amino-transferases de *Arabidopsis* (TAA) e das YUC flavinas mono-oxigenases (YUC) para produção de AIA a partir de Trp (Figura 4), Yue *et al.* (2014) fizeram estudos filogenéticos para verificar as origens dessas duas enzimas. Segundo esses autores, os genes codificantes das TAAs estão relacionados a homólogos em eucariotos fotossintéticos secundários (ex.: coanoflagelados e seus parentes próximos apusozoários e ichthyosporos) que não produzem AIA, que é consistente com o fato deles não possuírem genes homólogos da YUC relacionada com a síntese de AIA. Por outro lado, Yue *et al.* (2014) encontraram homólogos da família de genes YUC em bactérias e em alguns eucariontes, entre eles algas marrons e algas verdes clorofiladas. Entretanto, os mesmos autores não encontraram homólogos de YUC em carófitas, que é o grupo de algas verdes mais próximo das plantas terrestres. A explicação mais aceitável para esse fato, segundo estes pesquisadores, é de que as YUCs necessárias para a produção de AIA foram adquiridas pelo ancestral comum das plantas terrestres por transferência horizontal de genes a partir de bactérias. Assim, para esses autores, a origem mosaica dessa via de produção indica que essa via é uma inovação das plantas terrestres, assim como a importância da transmissão lateral de genes na geração e optimização de novidades evolutivas nos eucariotos.

Após entrarem nas células, as auxinas desencadeiam um conjunto de respostas celulares mediadas por um grupo de proteínas presentes no núcleo da célula vegetal. O receptor de auxina, *TRANSPORT INHIBITOR RESPONSE 1* (TIR1), interage com a proteína Culina (em inglês, *Cullin*), com a *Arabidopsis* SKP1 (ASK1/ASK2) uma proteína quinase tipo SKP1 (em inglês *S-phase kinase-associated protein*) e com a proteína com domínio em anel RBX1(sigla em inglês para *Ring Box 1*) para formar o complexo do tipo SCF (SKP – Culina – F-box) com atividade de ubiquitina ligase (Moon *et al.* 2004; Bishopp *et al.* 2006). Esse tipo de complexo transfere uma ubiquitina para a proteína alvo, o que marca esta para ser degradada pelo proteossomo 26S. O complexo SCF^{TIR1}, quando ligado à auxina, marca para degradação um conjunto de reguladores de transcrição dependentes de auxina (Aux/IAA). Esses reguladores de transcrição, na ausência de auxina, reprimem a atividade transcricional de FATORES DE RESPOSTA A AUXINA (*AUXIN RESPONSE FACTORS* – ARF, em inglês). Assim, na presença da auxina, a Aux/IAA é marcada para degradação pelo complexo SCF^{TIR1}, o que libera ARF para promover a transcrição dos genes associados à resposta à auxina (Figura 4) (Bishopp *et al.* 2006)

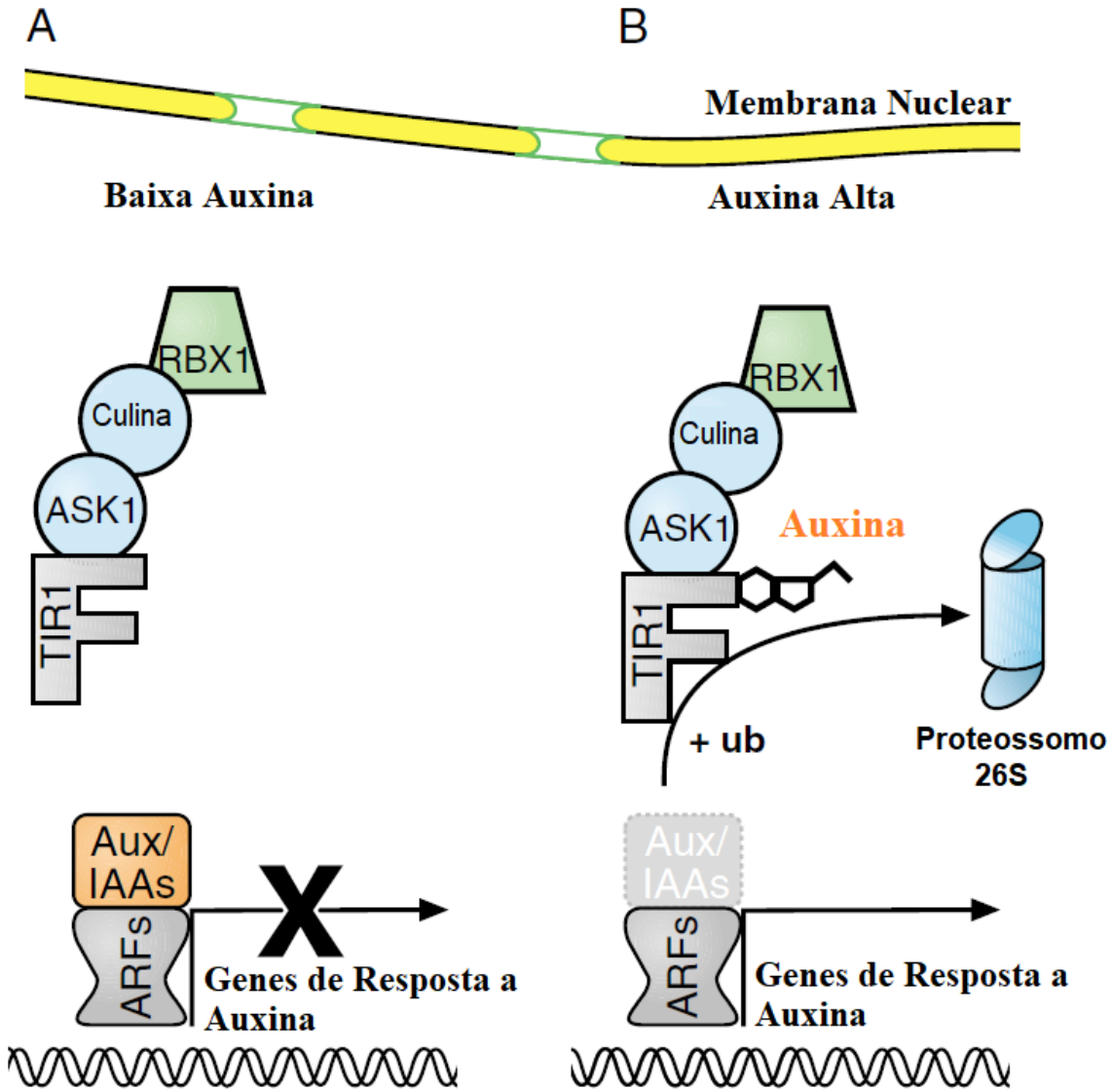


Figura 4. Mecanismo de ação das auxinas em plantas. (A) Em baixas concentrações de auxinas, a transcrição dos genes de resposta à auxina é bloqueada pela Aux/IAA. (B) Em altas concentrações de auxina, em *Arabidopsis*, esta se liga ao complexo SCF^{TIR1} pela TIR1. A ligação da auxina no complexo estimula a interação desse com a Aux/IAA. SCF^{TIR1} marca Aux/IAA para degradação por meio da ubiquitinação, que é feita pelo proteossoma 26S. Essa destruição libera as ARFs de seus repressores. Adaptada de Bishopp *et al.* 2006.

1.5.Transcritoma e o sequenciamento do RNA (RNA–seq):

Transcritoma é o conjunto formado por todos os transcritos de uma célula em um estágio desenvolvimental específico ou uma dada condição fisiológica. Estudar e entender o transcriptoma

é essencial para se compreender os elementos funcionais do genoma e seus constituintes, além de permitir estudar os estágios do desenvolvimento de dado organismo e aspectos relacionados a doenças. O estudo do transcritoma tem por objetivos (Wang *et al.* 2009):

- Catalogar os tipos de transcritos (RNA mensageiro, RNA não codante, pequenos RNAs, etc);
- Determinar a estrutura transcricional dos genes (por exemplo: sítios de início da transcrição, padrão de "splicing" e edição do RNA);
- Quantificar as mudanças nos níveis de expressão de cada transcrito durante o desenvolvimento ou sob determinada condição.

Dentre as tecnologias usadas para se estudar o transcritoma, a mais recente é a técnica de sequenciamento de RNA (RNA–seq). Essa técnica surgiu a partir do desenvolvimento de uma nova técnica de sequenciamento de DNA com alto rendimento. Nessa metodologia, uma população de RNA é convertida em uma biblioteca de fragmentos de cDNA com adaptadores ligados em ambas ou somente uma das extremidades do mesmo. Depois, cada fragmento, com ou sem amplificação prévia, é sequenciado e se obtém pequenas sequências. O sequenciamento é feito a partir de uma das extremidades (*single-end sequencing*) ou de ambas (*pair-end sequencing*). As leituras apresentam comprimento de 30 a 400 pares de bases (pb), dependendo da tecnologia de sequenciamento de DNA usada. As vantagens dessa metodologia frente às anteriores são (Wang *et al.* 2009):

- Diferentemente das metodologias baseadas em hibridização, RNA-seq não necessita de prévio sequenciamento do genoma do organismo sob estudo;
- RNA-seq pode revelar de forma precisa onde se localiza os limites dos transcritos com a resolução de 1 pb;
- RNA-seq também pode ajudar a identificar variações na sequência (por exemplo, polimorfismo de nucleotídeo único ou, em inglês, *single nucleotide polymorphism* – SNP) nas regiões transcritas.

Recentemente, com o auxílio dessa técnica, se começou a estudar a expressão de organismos que interagem no ambiente, em relações como o comensalismo, mutualismo e hospedeiro-patógeno (Wolf *et al.* 2018). A esse tipo de experimento que analisa, de forma simultânea, o transcritoma de ambos os organismos durante a interação e como estes influenciam a expressão um do outro, se deu o nome de Dual RNA-seq (Westermann *et al.* 2012; Westermann *et al.* 2016; Westermann and Vogel, 2018). Os experimentos de Dual RNA–seq são compostos das seguintes etapas (Wolf *et al.* 2018):

I. Execução do experimento onde dois ou mais organismos interagem;
II. Extração do RNA total, que contém RNA de ambos os organismos sob estudo;
III. Sequenciamento das bibliotecas de cDNAs geradas a partir das amostras de RNA total obtidas;
IV. Controle de qualidade e trimagem das sequências;
V. Mapeamento das sequências contra os genomas de referência;
VI. Determinação da quantidade de "reads" por gene;
VII. Identificação dos genes que estão sendo diferencialmente expressos nas condições do experimento.

As etapas de I a III correspondem às etapas de construção das bibliotecas e as demais correspondem ao processamento e análise dos dados obtidos. A etapa mais complicada neste tipo de análise é a separação *in silico* das sequências de cada organismo. Essa separação ocorre durante o mapeamento e pode ser feita usando duas técnicas de análise diferentes: sequencial e combinada (Wolf *et al.* 2018; Westermann *et al.* 2012; Westermann *et al.* 2016; Westermann and Vogel, 2018; Aprianto *et al.* 2016; Daly *et al.* 2017; Camilios-Neto *et al.* 2014; Pankievicz *et al.* 2016; Verwaaijen *et al.* 2017).

A análise sequencial consiste em mapear a biblioteca de cDNA primeiro contra o genoma de referência de um dos organismos. As sequências que não mapearem no primeiro genoma são consideradas como pertencentes ao outro organismo. Assim, essas sequências são mapeadas contra o outro genoma de referência (Verwaaijen *et al.* 2017; Packard *et al.* 2017; Camilios-Neto *et al.* 2014). Variantes desse método são a análise paralela e a combinada. Na primeira, a biblioteca é mapeada contra ambos os genomas de referência (Westermann *et al.* 2012; Westermann *et al.* 2016; Westermann and Vogel, 2018). Na análise combinada, as bibliotecas são mapeadas contra um genoma de referência quimérico, contendo ambos os genomas (Aprianto *et al.* 2016). Em ambos os casos, as sequências que alinham em ambos os genomas são eliminadas, pois não há consenso sobre o uso ou descarte dessas sequências.

Outro fato a se considerar quando da análise das bibliotecas oriundas de experimentos de Dual RNA–seq é como serão tratadas as sequências que alinham em mais de um local. Essas sequências podem ser de dois tipos: as que alinham em mais de um local no genoma e as que alinham em mais de um genoma. As sequências que podem alinhar em mais de um local no genoma de referência (Conesa *et al.* 2016) ou em ambos os genomas de referência (mapeamento cruzado) (Westermann *et al.* 2012; Westermann and Vogel, 2018) são chamadas de "multireads". De acordo com Conesa e colaboradores (2016), sequências que podem alinhar em mais de um local no

genoma (“multireads” genômicos) são principalmente devidas a sequências repetidas ou domínios compartilhados por genes parálogos no genoma. Essas sequências correspondem a uma fração significante das sequências mapeadas e, por isso, não devem ser descartadas (Conesa *et al.* 2016), sob pena de se perder informação sobre a expressão gênica, caso sejam eliminadas. Com relação às sequências que podem alinhar em mais de um genoma, deve-se levar em consideração o fato que procariotos e eucariotos que vivem no mesmo ambiente e interagem entre si podem trocar sequências de DNA através da Transferência Horizontal de Genes (HGT –*Horizontal Gene Transfer*, em inglês). Esse fato deve ser levado em conta, uma vez que há evidências de que DNA pode ser transferido de bactérias para eucariotos (Husnik and McCutcheon, 2017; Quispe-Huamanquispe, Gheysen and Kreuze, 2017) e de plantas para bactérias (Pontiroli *et al.* 2009). Dessa forma, a existência de similaridades entre algumas das sequências de ambos os genomas é esperada. Essas similaridades podem dificultar a atribuição correta da sequência ao genoma de referência. Por essa razão, segundo Wolf e colaboradores (2018), no caso de bibliotecas formadas pelos transcritos de dois ou mais organismos a análise combinada é a mais recomendada, pois permite à ferramenta de mapeamento analisar em qual genoma de referência a sequência alinha melhor.

Assim, dada a importância da cultura do milho e da necessidade de se compreender melhor como ocorre a interação entre este e a PGPB *Azospirillum barsilense* se realizou o estudo apresentado nesta tese de doutoramento. Esse estudo foi desenvolvido para entender melhor como o ácido indol – 3 – acético produzido por esta bactéria pode alterar a expressão gênica da planta.

2. Objetivo Geral

Analisar os efeitos na expressão gênica durante a interação de milho (*Zea mays*) com *Azospirillum brasilense* durante a inibição da produção do Ácido Indol-3-Acético pela planta.

2.1. Objetivos Específicos

1) Desenvolver de uma metodologia para a análise de bibliotecas de RNA compostas por dois ou mais organismos.
2) Avaliar a alteração da expressão gênica de plântulas de milho e da bacteria *Azospirillum brasilense* em condições de inoculação planta-bactéria e na presença de um inibidor da via de síntese de Ácido-Indol-3-Acético (Yucasin) em amostras de raízes.

3. Material, Métodos e Resultados:

Os materiais e métodos, e resultados serão apresentados no formato de artigo nos próximos dois tópicos. O artigo 1 foi submetido à revista *Genetics and Molecular Biology*, dessa forma foi escrito e apresentado de acordo com as normas desta revista. O artigo 2 será submetido à revista *Molecular Plant – Microbe Interactions* e por isso foi escrito e apresentado de acordo com as normas desta revista.

4. Artigo 1: The combined analysis as the best strategy for Dual RNA-Seq mapping

The combined analysis as the best strategy for Dual RNA-Seq mapping

Eliandro Espindula[1], Edilena Reis Sperb[1], Evelise Bach[1], Luciane Maria Pereira Passaglia[1,*]

[1]Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

* Corresponding author: LMP Passaglia, IB-UFRGS, Av. Bento Gonçalves 9500, prédio 43312, sala 207b, 91501-970 Porto Alegre, Brazil. Fax: +5551 3308 9816; e-mail: luciane.passaglia@ufrgs.br

Artigo submetido à revista Genetics and Molecular Biology (em fase de revisão)

Abstract

In Dual RNA-Seq experiments the simultaneous extraction of RNA and analysis of gene expression data from both interacting organisms could be a challenge. One alternative is separating the reads during *in silico* data analysis. There are two main mapping methods used: sequential and combined. Here we present a combined approach in which the libraries were aligned to a concatenated genome to sort the reads before mapping them to the respective annotated genomes. A comparison of this method with the sequential analysis was performed. Two RNA-Seq libraries available in public databases consisting of a eukaryotic (*Zea mays*) and a prokaryotic (*Herbaspirillum seropediceae*) organisms were mixed to simulate a Dual RNA-Seq experiment. Libraries from real Dual RNA-Seq experiments were also used. The sequential analysis consistently attributed more reads to the first reference genome used in the analysis (due to cross-mapping) than the combined approach. More importantly, the combined analysis resulted in lower numbers of cross-mapped reads. Our results highlight the necessity of combining the reference genomes to sort reads previously to the counting step to avoid losing information in Dual RNA-Seq experiments. Since most studies first map the RNA-Seq libraries to the eukaryotic genome, much prokaryotic information has probably been lost.

Keywords

Dual RNA-Seq; sequential analysis; combined analysis; mapping strategies.

Introduction

Organisms modulate gene expression to establish many interactions, from pathogenic to beneficial relationships (Wolf *et al.* 2018). There is a myriad of eukaryotic-prokaryotic interaction systems being studied, mainly focusing on pathogens and host gene expression responses, and pathogen-associated molecular patterns (PAMPs) (Westermann *et al.* 2012). Besides that, another successful molecular interaction being widely studied is the relationship between plants and beneficial plant growth promoting bacteria (PGPB), which finds application in the understanding of agricultural inoculants (Balsanelli *et al.* 2016; Bruto *et al.* 2014; Camilios-Neto *et al.* 2014).

Changes in gene expression or transcriptomes were first studied by microarray experiments. However, the simultaneous evaluation of both interacting organism transcriptomes by microarray presents some technical limitations as cross-hybridization, low sensitivity, and the cost of developing new probes to each new experiment (Westermann *et al.* 2012). Thus, these transcriptome studies were mostly performed focusing on only one of the interacting organisms (Barret *et al.* 2009; Mela *et al.* 2011). The RNA sequencing methodology (RNA-Seq) was developed to overcome microarray disadvantages and constitutes a promising approach for the parallel study of both interacting organisms, which was called Dual RNA-Seq (Westermann *et al.* 2012). In the beginning, this technique also presented restrictions related to cost and a significant amount of data management, which is being surpassed by the advent of new sequencing methodologies and bioinformatic tools. However, many RNA-Seq experiments still focused in only one organism of the interaction (Boscari *et al.* 2012; Hegedűs *et al.* 2009; Pankievicz *et al.* 2016; Verwaaijen *et al.* 2017), whereas others assessed the transcriptome of both interacting organisms (Aprianto *et al.* 2016; Choi *et al.* 2014; Reeder *et al.* 2017;

Westermann *et al.* 2012; Westermann *et al.* 2016; Westermann and Vogel, 2018; Wolf *et al.* 2018).

To perform a Dual RNA-Seq, steps of RNA isolation from both organisms, rRNA depletion, and cDNA library construction were adapted from the ones applied to simple RNA-Seq experiments (Westermann *et al.* 2012). To analyze Dual RNA-Seq data, there are two approaches to choose: sequential or combined analysis (Wolf *et al.* 2018). As the names infer, the former consists of the sequential analysis of the libraries against the reference genomes, one after the other (Camilios-Neto *et al.* 2014). In this approach, reads that fail to map to the first chosen reference genome are assumed to belong to the second genome. Therefore, these unmapped reads are the only ones used to map to the second genome (Packard *et al.* 2017; Verwaaijen *et al.* 2017). Variants of this methodology are the parallel and the combined analyses. In the first, the libraries are aligned simultaneously against both reference genomes (Westermann *et al.* 2016; Westermann and Vogel, 2018), and, in the last one, to a chimeric reference genome by concatenating the reference genomes (Aprianto *et al.* 2016). In both cases, all reads that aligned equally well to both genomes or have low alignment accuracy are removed.

Even though both methodologies described above are used to analyze Dual RNA-Seq data, they apparently are simple adaptations from the RNA-Seq methodologies that analyze one transcriptome at the time (Förstner *et al.* 2014; Westermann *et al.* 2016) and it seems critical to compare and evaluate which is the best choice for Dual RNA-Seq experiments. It is also worth considering that there is no consensus about the use or not of sequences that can align in more than one genome. In spite of the simultaneous read mapping being suggested in 2012 (Westermann *et al.* 2012), most of the Dual RNA-Seq works still opt to use the sequential approach (Kovalchuk *et al.* 2019; LaMonte *et al.* 2019; Mateus *et al.* 2019; Montoya *et al.* 2019; Mutha *et al.* 2019).

Here we present a mapping strategy for the combined analysis that consists of: i) aligning the Dual RNA-Seq libraries against a single file containing both reference genomes; ii) after this first mapping procedure, the reads attributed to each genome are extracted and saved into separated files; iii) these files are then used as individual libraries for the counting step using the respective annotated genome. Besides that, we present comparisons of this methodology to the sequential analysis to emphasize the importance of carefully choosing the mapping strategies for Dual RNA-Seq analysis. We test our approach using RNA-Seq libraries from different interaction systems. In two of them, we used data available in public databases, whereas in another analysis, the RNA-Seq libraries were part of an experiment performed in our laboratory that aimed to study the interaction between *Glycine max* roots and the bacterium *Bradyrhizobium elkanii.*

Material and Methods

RNA-Seq libraries and reference genomes

In order to test the combined analysis, we used RNA-Seq libraries available in public databases.

Firstly, data from two independent works were used to simulate a Dual RNA-Seq library. These data consisted in: NT-2 (replicate b) library from the bacterium *Herbaspirillum seropedicae* SmR1, available in the ArrayExpress database under the accession number E-MTAB-2842 (Bonato *et al.* 2016); and four mRNA libraries isolated from the central portion of the starchy endosperm of *Zea mays* (maize) cv. B73 six days after pollination, available in the NCBI database under the accession number SRP043224 (Thakare *et al.* 2014). Maize libraries were merged into a single file.

Our second dataset was comprised of dual RNA-Seq paired-end data from Lanubile *et al.* (2014), who investigated maize root genes involved in the defensive response to the infection caused by the fungus *Fusarium verticillioides*. Libraries of the

biological replicates of the susceptible maize variety CO354 inoculated with *F. verticillioides* were obtained from NCBI, accession numbers SRR1186869, SRR1186870, and SRR1186871 (Lanubile *et al.* 2014).

Finally, we used Dual RNA-Seq single-end data obtained in our laboratory. The experiment was designed to evaluate the interaction between two varieties of soybean (*Glycine max*, EMBRAPA 48 and BR 16) with the bacterium *Bradyrhizobium elkanii* strain SEMIA 587. Libraries were obtained as described below and were deposited at NCBI under the accession numbers: SRR7206486: BR16, replicate I; SRR7206485: BR16, replicate II; SRR7206490: EMBRAPA 48, replicate I; SRR7206489: EMBRAPA 48, replicate II.

Reference genomes of *B. elkanii* USDA76 (GCF_000379145.1), *G. max* (GCF_000004515.5), *H. seropedicae* Z67 (GCF_001040945.1), *Fusarium verticillioides* (GCF_000149555.1) and *Z. mays* cv. B73 (GCF_000005015.2) and their respective annotations were obtained from NCBI.

Data analysis

The CLC Genomics Workbench 8.0 (CLC – Bio; QIAGEN) toolkit was used to perform the trimming, mapping, and counting steps. The "Trimming" tool was used to trim reads smaller than 20 nucleotides from the RNA-Seq libraries, according to the program default settings for quality control. The "Convert to Tracks" tool was applied to the reference genomes to correctly associate them to the respective annotations.

Trimmed *H. seropedicae* and *Z. mays* RNA-Seq libraries were aligned to their respective reference genomes to eliminate possible contaminant reads, using the "Map to a Reference" tool with the parameters set to 0.8 of minimum length fraction and 0.8 of minimum similarity fraction. This procedure was called direct mapping and the libraries were called filtered libraries (Figure S1A). Both filtered libraries (from *Herbaspirillum*

138 or maize) were exported as separate fastq files, which were further merged into a single
139 file to form a Chimera Library to simulate a Dual-RNA-Seq experiment (Figure S1B).

140 We considered cross-mappings, the number of reads that belonged to one
141 organisms' transcriptome that mapped to the other organisms' genome. To check for
142 cross-mapping, each RNA-Seq filtered library was aligned to the reference genome of the
143 other organism (Figure S2A). Both cross-mapping and contamination checking steps
144 were useful to further evaluate our results. The *H. seropedicae* and *Z. mays* reference
145 genomes were also merged into a single file (Combined Reference), and each RNA-Seq
146 filtered library from *Herbaspirillum* and maize was aligned to the Combined Reference
147 file (Figure S2B).

148 The Chimera Library was used for the sequential and combined analyses and
149 mapping was done with the "Map to a reference" tool of CLC's program. The first
150 sequential analysis was performed aligning reads to the maize reference genome to
151 generate the first set of data (Eukaryote first- Figure 1A). Afterward, the exact opposite
152 was performed, and the reads were mapped against the bacterium reference genome first
153 to produce the second set of data (Prokaryote first- Figure 1B). In the combined analysis,
154 we aligned the Chimera Library to the Combined Reference file to sort out the sequences
155 belonging to one or another genome (Figure 1C). After sorting the sequences, those
156 attributed to each genome were extracted and exported as separate fastq files. Files were
157 imported back to CLC to count the reads of each library as described below.

158 Reads from RNA-Seq libraries of *Z. mays*, *H. seropedicae,* and from the Chimera
159 library that aligned to tRNA, rRNA, and to CDS (coding DNA sequence) loci were
160 counted using the CLC's tool "RNAseq" with the parameters set to 0.8 of minimum length
161 fraction and 0.8 of minimum similarity fraction, not mapping to intergenic regions, and
162 allowing a maximum of 5 hits (Figure 1).

In order to compare the results observed for the Chimera Library, the dual RNA-Seq libraries obtained in the soybean/*Bradyrhizobium* and maize/*Fusarium* experiments (Lanubile *et al.* 2014) were also analyzed using the sequential and combined approaches. In the maize/*Fusarium* experiment, some reads were mapped as broken pairs. Although these reads could align independently, none of the possible placements of the pair satisfied the pairing criteria. These reads were then treated as independent and marked as broken pairs. As these reads satisfied the mapping criteria, they were maintained in the following steps of the analysis.

Soybean varieties, bacterial strain, inoculation, growth, and experimental conditions

Soybean (*Glycine max)* plants of the contrasting genotypes EMBRAPA 48 and BR 16 (Oya *et al.* 2004) were grown under controlled temperature (26 ± 4ºC), luminosity (~ 600 μmol m^{-2} s^{-1}), and photoperiod (18/6 h light/dark). Cultivation was carried out in magenta boxes sealed in the root system, under a hydroponic system. Nutrients were supplied through the Hoagland's nutritive solution ½ strength (Hoagland and Arnon, 1938), which was replaced every three days. The nutritive solution was modified lacking nitrogen to stimulate nodulation. Soybean seeds were surface-sterilized by washing them three times with autoclaved ultrapure water, followed by soaking them in 70% ethanol for 3 min, and by a solution of 2% sodium hypochlorite and 2.5% Tween 20 for 30 min. Seeds were then washed three times with sterile distilled water by gentle shaking (Faleiro *et al.* 2013). All solutions and materials used were sterilized at 120ºC for 30 min. When the V2-V3 stage (Fehr *et al.* 1971) was reached, seedlings were inoculated with the symbiotic bacterium *Bradyrhizobium elkanii* SEMIA 587. *B. elkanii* was cultivated in yeast-mannitol liquid medium (Somasegaran and Hoben, 1994) in an orbital shaker (28°C, 120 rpm). When cultures reached an OD_{600} of 0.6, they were collected and centrifuged for 10 min at 10,000 x *g* at 4°C. The resulting pellets were washed twice with

sterile 0.85% NaCl solution, suspended in the same solution, and then diluted to obtain the inoculation solution at a concentration of approximately 10^8 CFU/mL (colony forming units).

Inoculation of the roots was performed by submerging them into the inoculant solution for 60 s. Inoculated roots were immediately frozen in liquid nitrogen and cryopreserved at -80°C for subsequent RNA isolation. Two biological replicates composed of pooled root seedlings from five plants were used for each genotype, resulting in four composed samples for further RNA isolation.

RNA isolation, mRNA enrichment, cDNA synthesis, and sequencing

Total RNA isolation of *G. max* root seedlings inoculated with *B. elkanii* was done using TRIzol (Invitrogen) reagent. The integrity of RNA was verified on 1.5% agarose gel. Concentration and purity were determined by spectrophotometry at 260 nm and 280 nm (Jahn *et al.* 2008) measured in Nanodrop LITE spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). RNA samples were subjected to a purification step using PureLink™ RNA Micro kit (Ambion), treated with DNaseI (Invitrogen) and then rRNA was depleted using the RiboMinus™ Plant Kit for RNA-Seq (Invitrogen). The cDNA libraries were constructed using the Ion total RNA-Seq kit v2 for Whole Transcriptome Library. All RNA quantification and quality evaluation were performed at the Bioanalyzer™ - Agilent 2100 instrument. Each cDNA library obtained was sequenced using the Ion PI Template OT2 200 Kit v3 and the Ion PI Sequencing 200 Kit v3 at the IonTorrent® platform (Thermo Fisher Scientific, Wilmington, DE, USA). All kits and reagents were used according to manufacturer's instructions.

The presence of the bacterium in plant roots was subsequently determined by the detection of its 16S rRNA gene sequences in the transcriptome library.

Results

Data analysis using independent RNA-Seq libraries

Before starting the analysis, trimmed RNA-Seq libraries from *Z. mays* and *H. seropedicae* were filtered by direct mapping to each genome to avoid potential contamination sequences (Figure S1A). After filtering, the *H. seropedicae* RNA-Seq library presented approximately 8.8 million reads, while the *Z. mays* RNA-Seq library presented approximately 22 million reads. The Chimera Library, which simulates a Dual RNA-Seq experiment, was constructed joining these two libraries (Figure S1B) and presented approximately 31 million reads (Table 1).

Cross-mappings were determined by the number of reads from one organism RNA-Seq library that could be attributed to the other organism reference genome (Figure S2A). Interestingly, approximately 2 million reads from the *H. seropedicae* RNA-Seq library aligned to the *Z. mays* genome, while 7,659 reads from the *Z. mays* RNA-Seq library mapped to the *H. seropedicae* genome (Table 1). On the other hand, when we mapped the individual RNA-Seq libraries to the Combined Reference file (Figure S2B), which was constructed by concatenating *H. seropedicae* and *Z. mays* genomes (Figure 1), more surprising results were obtained. When the *H. seropedicae* RNA-Seq library was aligned to the Combined Reference file, approximately 8.7 million reads were attributed to *H. seropedicae* genome and 86,366 reads to *Z. mays* genome; whereas when *Z. mays* RNA-Seq library was mapped to the Combined Reference file, 394 reads were attributed to *H. seropedicae* genome and approximately 22 million reads to *Z. mays* genome (Table 1). These results showed that even in the presence of both reference genomes some reads still mapped incorrectly, although the numbers of reads incorrectly aligned were much smaller than the numbers of cross-mapping reads obtained when one RNA-Seq library was aligned to the genome of the other organism.

After estimating cross-mappings, we evaluated both the sequential and the combined approach of Dual RNA-Seq analysis. The sequential analysis consisted of aligning the Chimera Library to one reference genome before the other. Reads that aligned to the first genome constituted this organism's library. Unmapped reads are then mapped to the second reference genome. All reads that aligned to the second genome comprised this organism's library. We first mapped the Chimera Library to *Z. mays* reference genome, and then the unmapped reads were mapped to the *H. seropedicae* reference genome (Figure 1A). Approximately 6 million reads were attributed to *H. seropedicae* genome, and approximately 24 million reads were attributed to *Z. mays* genome (Table 1 - Eukaryote 1st). When we did the opposite and first mapped the Chimera Library to the *H. seropedicae* reference genome (Figure 1B), approximately 8 million reads were attributed to *H. seropedicae* genome and approximately 22 million reads to *Z. mays* genome (Table 1 – Prokaryote 1st).

Finally, we performed the combined analysis that consists of aligning the RNA-Seq library to a file containing a combination of reference genomes (Combined Reference). We mapped the Chimera Library to the Combined Reference file, and this alignment approach attributed approximately 8.7 million reads to *H. seropedicae* genome and approximately 22 million reads to *Z. mays* genome (Table 1 – Combined analysis, Figure 1C). After the mapping procedure, reads attributed to each genome were extracted, saved into separated files (Figure S3), and used as individual libraries for the counting step (Figure 1C). All reads that aligned to *H. seropedicae* or *Z. mays* genomes were counted using the corresponding reference genome and its respective annotations (Tables 2 and 3). In all counts we observed unmapped reads. This is likely due to the parameters chosen for counting the reads in the CLC's "RNAseq" tool, as reads that mapped in more than five loci or mapped in intergenic regions were excluded.

The counting of reads that aligned to the respective genome in the direct, sequential, and combined analysis showed interesting results (Table 2). In the direct analysis, for the *H. seropedicae* RNA-Seq library, 235,428 reads were attributed to tRNA, 5,649,357 reads to rRNA, and 2,307,178 reads to CDS loci using the *H. seropedicae* genome, 616,765 remained unmapped, while for *Z. mays* RNA-Seq library, we counted 1,692 tRNA reads, 3,003 rRNA reads, and 21,051,646 CDS loci reads using the *Z. mays* genome, 1,144,534 remained unmapped (Table 2 – Direct mapping). In the sequential analysis, when we first mapped the Chimera Library to the *Z. mays* genome, we counted using the genome of *H. seropedicae* 193,088 tRNA reads, 4,263,732 rRNA reads, 1,616,947 CDS loci reads, and 482,355 reads remained unmapped, while when using the genome of *Z. mays*, we counted 13,175 tRNA reads, 403,522 rRNA, 22,119,380 CDS loci reads, and 1,917,404 reads remained unmapped (Table 2 – Eukaryote 1st). On the other hand, when we first mapped the Chimera Library to the *H. seropedicae* genome and also counted with *H. seropedicae* files, 235,430 reads were attributed to tRNA, 5,650,140 reads to rRNA, 2,313,367 to CDS loci, and 617,450 reads remained unmapped, while when counting using the *Z. mays* genome, 1,686 reads were attributed to tRNA, 2,255 reads to rRNA, 21,045,092 reads to CDS loci, and 1,144,183 reads remained unmapped (Table 2 – Prokaryote 1st). Finally, when we counted the reads that mapped to each reference genome using the combined analysis and counted using *H. seropedicae* files, 234,656 reads were attributed to tRNA, 5,618,104 reads to rRNA, 2,258,318 reads to CDS loci, and 608,264 reads remained unmapped, while when counting with the *Z. mays* genome, 1,739 reads were attributed to RNA, 11,372 reads to rRNA, 21,104,107 reads to CDS loci, and 1,169,608 reads remained unmapped (Table 2 – Combined analysis).

In order to investigate the reads that were incorrectly aligned (cross-mapped reads), those reads were also counted using both the respective and the incorrect reference

genomes (Figure S2, Tables 3, S1 and S2). Although several reads incorrectly mapped to rRNA and tRNA loci, the most important result was that about one million *H. seropedicae* reads were incorrectly attributed to almost 10,000 *Z. mays* CDS. A similar situation, although with minor effects, was observed for the *Z. mays* library, where 6,189 reads were incorrectly attributed to 65 *H. seropedicae* CDS. In these cases, only CDS that received at least ten reads assigned to them were considered. When the combined reference file was used, the numbers of reads incorrectly mapped decreased significantly, in particular for the *H. seropedicae* genome. In this case, 52,530 reads from *H. seropedicae* were incorrectly attributed to 575 *Z. mays* CDS.

Tables S1 and S2 present the top 20 most counted loci among the cross-mapped reads. Table S1 shows the loci where the reads should be aligned in the correct genome, while Table S2 shows the loci where the reads aligned in the incorrect genome. It is interesting to note that most of the incorrectly mapped reads corresponded to genes that code for proteins with different functions, such as kinases, phosphatases, and ribosomal proteins. Several genes coding for hypothetical or uncharacterized proteins were also identified.

Analyzes of Dual RNA-Seq libraries experimentally obtained

The combined Dual RNA-Seq analysis was also applied to RNA-Seq libraries obtained from two experiments, one performed in our laboratory and another carried out by Lanubile *et al.* (2014). The first one aimed to evaluate the interaction of two varieties of *G. max* with the symbiotic bacterium *B. elkanii.* The presence of the bacterium in the plant's roots was confirmed by the detection of its 16S rRNA gene sequences in the RNA-Seq libraries (data not shown). RNA-Seq libraries obtained from both organisms showed enough quality and coverage to perform gene expression analysis (Table S3). After the trimming procedure, RNA-Seq libraries from the *G. max* - *B. elkanii* experiment

313 presented approximately 5 to 9 million reads (Table S3). The Dual RNA-Seq alignment
314 strategies showed that numbers attributed to the eukaryotic genome roughly did not vary
315 among sequential or combined analyses, regardless of the soybean variety used (Figure
316 2A, Table S3). However, some variation was observed in the number of reads mapped to
317 the prokaryotic genome depending on the mapping approach. When RNA-Seq reads were
318 first aligned to the eukaryotic genome (Table S3 - Sequential analysis- Eukaryote 1st), the
319 number of reads attributed to the bacterium was less than 2% of the total amount of reads
320 (Figure 2B) for both soybean varieties. However, when the opposite analysis was
321 performed (Table S3 - Sequential analysis- Prokaryote 1st), the number of reads aligned
322 to the prokaryote genome increased significantly, reaching more than 3% of the total
323 amount of reads mapped in both samples (Figure 2B), also for both soybean varieties.
324 Using the combined analysis, when reads were aligned to both genomes at the same time,
325 intermediary numbers of reads were attributed to the prokaryote genome, regardless of
326 the soybean variety used (Table S3 – Combined analysis, Figure 2B). The average number
327 of reads attributed to the prokaryote in the combined analysis was not significantly
328 different from the average number attributed at the Sequential Analysis – Eukaryote 1st.
329 Despite this fact, the results still indicated that probably some reads that mapped to the
330 first genome used in the sequential approach very likely belong to the second genome and
331 incorrectly mapped to the first genome because the second was not present in the analysis.
332 The second experiment used to evaluate the combined Dual RNA-Seq analysis
333 was performed by (Lanubile *et al.* 2014), who investigated maize roots gene expression
334 during *Fusarium verticillioides* infection. Although these authors investigated *Z. mays*
335 genes only, library preparation involves the isolation of mRNAs using poly(A)-tails,
336 which potentially included fungus mRNA. Thus, we chose this library as another example
337 of plant-microorganism interaction. After the trimming procedure, the libraries had from

74 to 83 million reads (Table S4). In this analysis, even though the numbers of reads attributed to both genomes varied according to the previous experiments (Table S4), the average number of reads attributed to each genome according to the methodology used were not significantly different, since the standard errors were substantial (Figure 3). Nevertheless, the combined analysis showed an intermediate amount of reads attributed to each genome in comparison with the number of reads observed in the sequential analyzes (Table S4, Figure 3), which was similar to the previous analyses.

Discussion

RNA sequencing methodologies are revolutionizing the way we study gene expression. Unlike microarrays, to perform an RNA-Seq analysis there is no need for previous knowledge about the organism. Another advantage of RNA-Seq is that it enables global gene expression analysis since it allows access to different populations of RNA sequences from the organism (Oshlack *et al.* 2010; Wang *et al.* 2009). In the last decade, this technique was used to assess gene expression of many organisms and it has recently started to be used to assess the transcriptomes of interacting organisms, called Dual RNA-Seq (Baddal *et al.* 2015; Camilios-Neto *et al.* 2014; Hayden *et al.* 2014; Pankievicz *et al.* 2016; Westermann *et al.* 2016).

Despite the difficulties in obtaining libraries containing RNAs from both interacting organisms, there are also problems in sorting the reads *in silico*. The sequential approach seems to be the most common mapping method chosen, and the order of the genomes used in the analysis is chosen according to study interests (Baddal *et al.* 2015, Camilios-Neto *et al.* 2014, LaMonte *et al.* 2019, Mateus *et al.* 2019, Montoya *et al.* 2019). Sometimes the reads of one of the interacting organisms are not considered for the study and are discarded from the analysis (Packard *et al.* 2017; Verwaaijen *et al.* 2017; Lanubile *et al.* 2014). Similarly, reads that aligned equally well to either genome or simply cross-

363 mapped are also sometimes discarded (Westermann *et al.* 2016; Westermann and Vogel,
364 2018; Baddal *et al.* 2015)

365 Here we used a Combined Analysis, which consists in using a Combined
366 Reference file formed by merging the reference genomes files of both organisms to *in*
367 *silico* sort the reads that align to each genome. Once identified, they were extracted and
368 saved in separated files (Figure S3). The libraries formed by the reads of each organism
369 were then counted using the corresponding reference genome with their own annotations.
370 To perform these analyses, we used the CLC's tools set with the parameters usually used
371 to map eukaryotic libraries (Camilios-Neto *et al.* 2014).

372 Before testing the combined approach, we determined the number of cross-
373 mapped reads between the two RNA-Seq libraries using the reference genome of the other
374 organism of the Combined Reference file. After aligning them, the reads that mapped to
375 the incorrect genome (cross-mapped reads) were counted using both the correct and
376 incorrect reference genome. This was done to identify the loci where these reads were
377 aligned in the incorrect genome and the loci where they should be assigned in the correct
378 one (cross-mapping; Tables 3, S1 and S2). Our results showed that the combined analysis
379 consistently assigned a lower number of reads to the incorrect organism due to cross-map,
380 allowing the program to better attribute the reads to its corresponding genome, leading to
381 a lower number of cross-mappings (Table 2 and 3).

382 After these cross-map evaluations, two sequential analyses were performed, and
383 the obtained results were compared with the results from the combined analysis. For both
384 sequential analyses, it was possible to notice that the first genome used on the mapping
385 step was always benefited. We observed that the first genome used to map the reads
386 received the full number of reads that could cross-map with the genome of the other
387 organism (Table 1). We also noticed that even though many of the cross-mapping reads

388 mapped to rRNA genes, a significant number of cross-mapping reads were attributed to
389 CDS loci in all methodological approaches. However, in the combined analysis, the loss
390 of reads due to cross-mapping was lower than in the sequential analysis (Table 3). Also
391 interesting was the fact that the *H. seropedicae* genome lost more reads for the *Z. mays*
392 genome due to cross-mapping than the other way around.

393 To compare the *in silico* data with real Dual RNA-Seq samples, libraries from two
394 different Dual RNA-Seq experiments were submitted to both sequential and the combined
395 approaches. In both experiments, the results obtained were similar and followed the
396 results from the *in silico* data, with the combined analysis showing intermediary values
397 when compared to values attributed by the sequential analyses (Tables S3 and S4). For
398 the *G. max* - *B. elkanii* experiment, the average amount of reads attributed to the combined
399 analysis was significantly different only concerning the Sequential – Prokaryote 1st data
400 (Figure 2B). Schurch *et al.* (2016) recommended that at least three biological replicates
401 must be used in order to detect genes being differentially expressed. Since the *G. max* -
402 *B. elkanii* experiment contained only two biological replicates we hypothesized that with
403 more biological replicates these two methodologies should present significant differences
404 concerning the number of reads attributed to each organism.

405 Another interesting fact was observed in the Lanubile *et al.* (2014) experiment.
406 When comparing the average amount of reads attributed to each genome, regardless of
407 the methodology used, no significant differences were observed (Figure 3). Analyzing
408 our results, it seems that paired-end sequencing was also useful to make the two
409 eukaryotic genomes more distinguishable and less prone to cross-mappings (Figure 3).
410 Therefore, one should consider using paired-end libraries allied with the combined
411 analysis in order to reduce the number of cross-mappings during Dual RNA-Seq
412 experiments.

Since we detected that a significant number of cross-mapping reads aligned to gene coding regions of the genomes, we can assume that this happened because the interacting organisms should have similar metabolic pathways or due to homologous sequences. Eliminating these reads from the libraries before counting them represents a problem because a considerable amount of transcriptional information will be lost. Therefore, all reads that can align to both genomes (with different degrees of similarity to each genome) will align to the first genome used in the sequential mapping approach. This might lead to an overestimation of the expressed genes of the first genome used in the sequential mapping method. Similarly, the expressed genes of the second genome might be underestimated. This problem seems to be more critical for those interested in the prokaryotic transcriptome. Prokaryotic RNA is always less abundant in libraries prepared from mixed sources (Westermann *et al.* 2012); therefore, techniques that underestimate their read counts should be avoided. The combined analysis seems to be more reasonable to avoid these under/overestimations.

Aprianto *et al.* (2016) suggested a Dual RNA-Seq approach in which they aligned the libraries to a chimeric genome. To create this genome, they concatenated the *Streptococcus pneumoniae* genome as an extra chromosome of *Homo sapiens* and adjusted the annotated genomes. All procedures were performed with command-line, that demands some bioinformatic knowledge and programming skills. Another objective of our work was to describe a way to analyze the Dual RNA-Seq libraries without the need for high computational skills. Therefore, to perform the proposed combined analysis, the CLC Workbench was used. This program is user-friendly since it works with a graphic interface and has several internal tutorials, which demands only basic bioinformatic skills. Thus, newcomers in the Dual RNA-Seq experiments can perform relatively complex analysis in an easier bioinformatics environment. Another aspect, and according to

Baruzzo *et al.* (2017), CLC Workbench, alongside with Novoalign and STAR, is one of the best aligners for eukaryotes in use nowadays, even when using the standard or improved setups.

A critical step during a Dual RNA-Seq experiment is to separate *in silico* the reads that align to each genome. Another reason to use CLC Genomics Workbench is that after performing the mapping step, the program results in a file containing a list showing in which particular reference the reads aligned. Based on this list, during a combined analysis, the researcher can easily select and extract all the reads that aligned to each reference genome and save them into separate files (Figure S3). As these files will only contain the reads of one organism, the counting step can be performed using the reference genome and annotations of the corresponding organism.

As a conclusion, with the present, work we were able to show that Dual RNA-Seq results vary according to the mapping strategy chosen and this could lead to misinterpretations of the interactions between organisms. Our results showed that the combined analysis allows a smaller loss of reads due to cross-mapping. This fact avoids the loss of relevant information to the first genome chosen in the mapping step when the sequential analysis is used. Since most studies first align the RNA-Seq libraries to the eukaryotic genome, much prokaryotic information is probably being lost. Thus, to fully comprehend gene expression and communication between interacting organisms, we suggest adopting the combined mapping analysis in Dual RNA-Seq experiments.

Acknowledgments

This work was supported by the Brazilian funding agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), and by Instituto Nacional de Ciência e Tecnologia da Fixação Biológica do Nitrogênio (INCT-FBN), Brazil.

Conflict of interest

The authors declare that they have no conflicts of interest related to the subject matter or materials discussed in this article.

Author contributions

E Espindula and ER Sperb conceived, designed, and performed the experiments. E Espindula, ER Sperb, E Bach, and LMP Passaglia analyzed and/or interpreted the data. LMP Passaglia contributed to reagents and materials. E Espindula, E Bach, and LMP Passaglia wrote the manuscript.

References

Aprianto R, Slager J, Holsappel S , Veening J-W (2016) Time-resolved dual RNA-seq reveals extensive rewiring of lung epithelial and pneumococcal transcriptomes during early infection. Genome Biol 17: 198.

Baddal B, Muzzi A, Censini S, Calogero RA, Torricelli G, Guidotti S, Taddei AR, Covacci A, Pizza M, Rappuoli R, Soriani M, Pezzicoli A (2015) Dual RNA-seq of *Nontypeable haemophilus* influenzae and host cell transcriptomes reveals novel insights into host-pathogen cross talk. mBio 6.

Balsanelli E, Tadra-Sfeir MZ, Faoro H, Pankievicz VCS, de Baura VA, Pedrosa FO, Souza EM, Dixon R, Monteiro RA (2016) Molecular adaptations of *Herbaspirillum seropedicae* during colonization of the maize rhizosphere. Environm Microbiol 18: 2343-2356.

Barret M, Frey-Klett P, Guillerm-Erckelboudt A-Y, Boutin M, Guernec G, Sarniguet A (2009) Effect of wheat roots infected with the pathogenic fungus *Gaeumannomyces graminis* var. tritici on gene expression of the biocontrol bacterium *Pseudomonas fluorescens* Pf29Arp. MPMI 22:1611-1623.

488 Bonato P, Batista MB, Camilios-Neto D, Pankievicz VCS, Tadra-Sfeir MZ, Monteiro
489 RA, Pedrosa FO, Souza EM, Chubatsu LS, Wassem R, Rigo LU (2016) RNA-seq
490 analyses reveal insights into the function of respiratory nitrate reductase of the
491 diazotroph *Herbaspirillum seropedicae*. Environ Microbiol 18: 2677-2688.
492 Boscari A, del Giudice J, Ferrarini A, Venturini L, Zaffini A-L, Delledonne M, Puppo A
493 (2013) Expression dynamics of the *Medicago truncatula* transcriptome during the
494 symbiotic interaction with *Sinorhizobium meliloti*: which role for nitric oxide?
495 Plant Physiol 161:425-39.
496 Bruto M, Prigent-Combaret C, Muller D, Moënne-Loccoz Y (2014) Analysis of genes
497 contributing to plant-beneficial functions in plant growth-promoting rhizobacteria
498 and related Proteobacteria. Sci Rep 4: 6261.
499 Camilios-Neto D, Bonato P, Wassem R, Tadra-Sfeir MZ, Brusamarello-Santos LCC,
500 Valdameri G, Donatti L, Faoro H, Weiss VA, Chubatsu LS, Pedrosa FO, Souza
501 EM (2014) Dual RNA-seq transcriptional analysis of wheat roots colonized by
502 *Azospirillum brasilense* reveals up-regulation of nutrient acquisition and cell
503 cycle genes. BMC Genomics 15: 378.
504 Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, Santarius T, Stebbings LA,
505 Leroy C, Edkins S, Hardy C, Teague JW, Menzies A, Goodhead I, Turner DJ,
506 Clee CM, Quail MA, Cox A, Brown C, Durbin R, Hurles ME, Edwards PAW,
507 Bignell GR, Stratton MR, Futreal PA (2008) Identification of somatically acquired
508 rearrangements in cancer using genome-wide massively parallel paired-end
509 sequencing. Nature Genet 40: 722.
510 Choi YJ, Aliota MT, Mayhew GF, Erickson SM, Christensen BM (2014) Dual RNA-seq
511 of parasite and host reveals gene expression dynamics during filarial worm–
512 mosquito interactions. PLoS Negl Trop Dis 8: e2905.

513 Faleiro AL, Pereira TP, Espindula E, Brod FCA., Arisi ACM (2013). Real time PCR
514 detection targeting *nifA* gene of plant growth promoting bacteria *Azospirillum*
515 *brasilense* strain FP2 in maize roots. Symbiosis 61: 125–133.
516 Fehr WR, Caviness CE, Burmood DT, Pennington JS (1971) Stage of development
517 descriptions for soybeans, *Glycine max* (L.) Merrill1. Crop Sci 11: 929-931.
518 Förstner KU, Vogel J, Sharma CM (2014) READemption—a tool for the computational
519 analysis of deep-sequencing–based transcriptome data. Bioinformatics 30: 3421-
520 3423.
521 Hayden KJ, Garbelotto M, Knaus BJ, Cronn RC, Rai H, Wright JW (2014) Dual RNA-
522 seq of the plant pathogen *Phytophthora ramorum* and its tanoak host. Tree Genet
523 Genomes 10: 489-502.
524 Hegedűs Z, Zakrzewska A, Ágoston VC, Ordas A, Rácz P, Mink M, Spaink HP, Meijer
525 AH (2009) Deep sequencing of the zebrafish transcriptome response to
526 mycobacterium infection. Mol Immunol 46: 2918-2930.
527 Hoagland DR, Arnon D (1938) The water culture method for growing plants without soil.
528 Circ Calif Agric Exp Sta 347(Revised 1950): 32
529 Jahn CE, Charkowski AO, Willis DK (2008) Evaluation of isolation methods and RNA
530 integrity for bacterial RNA quantitation. J Microbiol Methods 75: 318-324.
531 Lanubile A, Ferrarini A, Maschietto V, Delledonne M, Marocco A, Bellin D (2014)
532 Functional genomic analysis of constitutive and inducible defense responses to
533 *Fusarium verticillioides* infection in maize genotypes with contrasting ear rot
534 resistance. BMC Genomics 15(1): 710.
535 LaMonte, G. M., P. Orjuela-Sanchez, J. Calla, L. T. Wang, S. Li, J. Swann, A. N. Cowell,
536 B. Y. Zou, A. M. Abdel-Haleem Mohamed, Z. H. Villa Galarce, M. Moreno, C.
537 Tong Rios, J. M. Vinetz, N. Lewis & E. A. Winzeler (2019) Dual RNA-seq

identifies human mucosal immunity protein Mucin-13 as a hallmark of *Plasmodium* exoerythrocytic infection. Nat Commun 10**,** 488.

Mateus, I. D., F. G. Masclaux, C. Aletti, E. C. Rojas, R. Savary, C. Dupuis & I. R. Sanders (2019) Dual RNA-seq reveals large-scale non-conserved genotype × genotype-specific genetic reprograming and molecular crosstalk in the mycorrhizal symbiosis. ISME J 13**,** 1226-1238.

Mela F, Fritsche K, de Boer W, van Veen JA, de Graaff LH, van den Berg M, Leveau JHJ (2011) Dual transcriptional profiling of a bacterial/fungal confrontation: *Collimonas fungivorans* versus *Aspergillus niger*. ISME J 5: 1494-1504.

Montoya, D. J., P. Andrade, B. J. A. Silva, R. M. B. Teles, F. Ma, B. Bryson, S. Sadanand, T. Noel, J. Lu, E. Sarno, K. B. Arnvig, D. Young, R. Lahiri, D. L. Williams, S. Fortune, BR. Bloom, M. Pellegrini, RL. Modlin (2019) Dual RNA-Seq of Human Leprosy Lesions Identifies Bacterial Determinants Linked to Host Immune Response. Cell Rep 26**,** 3574-3585.e3.

Mutha, N. V. R., W. K. Mohammed, N. Krasnogor, G. Y. A. Tan, W. Y. Wee, Y. Li, S. W. Choo, N. S. Jakubovics (2019) Transcriptional profiling of coaggregation interactions between *Streptococcus gordonii* and *Veillonella parvula* by Dual RNA-Seq. Sci Rep-UK, 9**,** 7664.

Oshlack A, Robinson MD, Young MD (2010) From RNA-seq reads to differential expression results. Genome Biol 11: 220.

PackardH, Burke AK, Jensen RV, Stevens AM (2017) Analysis of the in planta transcriptome expressed by the corn pathogen *Pantoea stewartii* subsp. stewartii via RNA-Seq. PeerJ 5: e3237.

Pankievicz VCS, Camilios-Neto D, Bonato P, Balsanelli E, Tadra-Sfeir MZ, Faoro H, Chubatsu LS, Donatti L, Wajnberg G, Passetti F, Monteiro RA, Pedrosa FO,

563 Souza EM (2016) RNA-seq transcriptional profiling of *Herbaspirillum*
564 *seropedicae* colonizing wheat (*Triticum aestivum*) roots. Plant Mol Biol 90: 589-
565 603.
566 Reeder SM, Palmer JM, Prokkola JM, Lilley TM, Reeder DM, Field KA (2017)
567 *Pseudogymnoascus destructans* transcriptome changes during white-nose
568 syndrome infections. Virulence 8(8): 1695-1707.
569 Somasegaran P, Hoben HJ (1994) Handbook for *Rhizobia*:Methods in Legume-
570 Rhizobium Technology. Springer-Verlag, New York, NY.
571 Thakare D, Yang R, Steffen JG, Zhan J, Wang D, Clark RM, Wang X, Yadegari R (2014)
572 RNA-Seq analysis of laser-capture microdissected cells of the developing central
573 starchy endosperm of maize. Genom Data 2: 242-245.
574 Verwaaijen B, Wibberg D, Kröber M, Winkler A, Zrenner R, Bednarz H, Niehaus K,
575 Grosch R, Pühler A, Schlüter A (2017) The *Rhizoctonia solani* AG1-IB (isolate
576 7/3/14) transcriptome during interaction with the host plant lettuce (*Lactuca sativa*
577 L.). Plos One 12: e0177278.
578 Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for
579 transcriptomics. Nat Rev Genet 10: 57-63.
580 Westermann AJ, Förstner KU, Amman F, Barquist L, Chao Y, Schulte LN, Müller L,
581 Reinhardt R, Stadler PF, Vogel J (2016) Dual RNA-seq unveils noncoding RNA
582 functions in host–pathogen interactions. Nature 529: 496-501.
583 Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-seq of pathogen and host. Nat
584 Rev Microbiol 10: 618-630.
585 Westermann AJ, Vogel J (2018) Host-Pathogen Transcriptomics by Dual RNA-Seq. In:
586 Arluison V, Valverde C (ed) Bacterial Regulatory RNA - Methods in Molecular
587 Biology. Humana Press, New York, pp. 59-75.

588 Wolf T, Philipp K, Brunke S, Linde J (2018) Two s company: studying interspecies
589 relationships with dual RNA-seq. Curr Opin Microbiol 42: 7–12.
590
591

592 **Figures and Tables legends**

593

594 **Figure 1.** Mapping strategies for Dual RNA-Seq analysis. (A) Sequential analysis
595 aligning libraries to the eukaryotic genome first- Eukaryote 1st; (B) Sequential analysis
596 aligning libraries to the prokaryotic genome first- Prokaryote 1st; (C) Combined analysis.

597

598 **Figure 2.** Percentage of reads mapped to (A) *Bradyrhizobium elkanii* or (B) *Glycine max*
599 depending on the methodology used in the *Glycine max – Bradyrhizobium elkanii*
600 experiment. Bars indicate twice the Standard Error. BR16 and ER48: soybean varieties
601 BR16 and Embrapa 48, respectively.

602

603 **Figure 3.** Percentage of reads mapped to (A) *Fusarium verticillioides* or (B) *Zea mays*
604 depending on the methodology used in the *Zea mays –Fusarium verticillioides*
605 experiment. Bars indicate twice the Standard Error.

606

607 **Table 1.** Library features and number of total reads attributed to the *Herbaspirillum*
608 *seropedicae* or *Zea mays* genomes according to the mapping approach. The analyses were
609 performed with the genomes without annotations. Cross-mapping: reads of each
610 individual library were aligned to the reference genome of the other organism. Sequential
611 Analysis: The library was first aligned to one reference genome; reads that fail to map to
612 the first genome were aligned to the other genome. Eukaryote 1st/Prokaryote 1st indicates
613 the first reference used. Combined analysis: libraries were mapped to a merged file
614 containing both reference genomes (Combined Reference).

615

Table 2. Number of reads mapped to tRNA, rRNA, and coding loci (CDS) according to the mapping methodology used.

Table 3. Comparison of the number of reads incorrectly mapped due to cross-mapping. Reads that incorrectly aligned to the reference genome were counted using the annotated genome indicated on the table. The unmapped reads are a result of the counting parameters that eliminate reads that aligned in more than five loci and in the intergenic regions. *CDS with at least 10 reads assigned to them. Exception made to the *Z. mays* library mapped to the Combined Reference, which refers to CDS with at least one read assigned to it.

Table 4. Number of reads mapped to tRNA, rRNA, and coding loci (CDS) according to the mapping methodology and experiment used. BR16 and ER48: soybean varieties BR16 and Embrapa 48, respectively. CO354: susceptible maize variety CO354 inoculated with *F. verticillioides* from Lanubile *et al.* (2014).

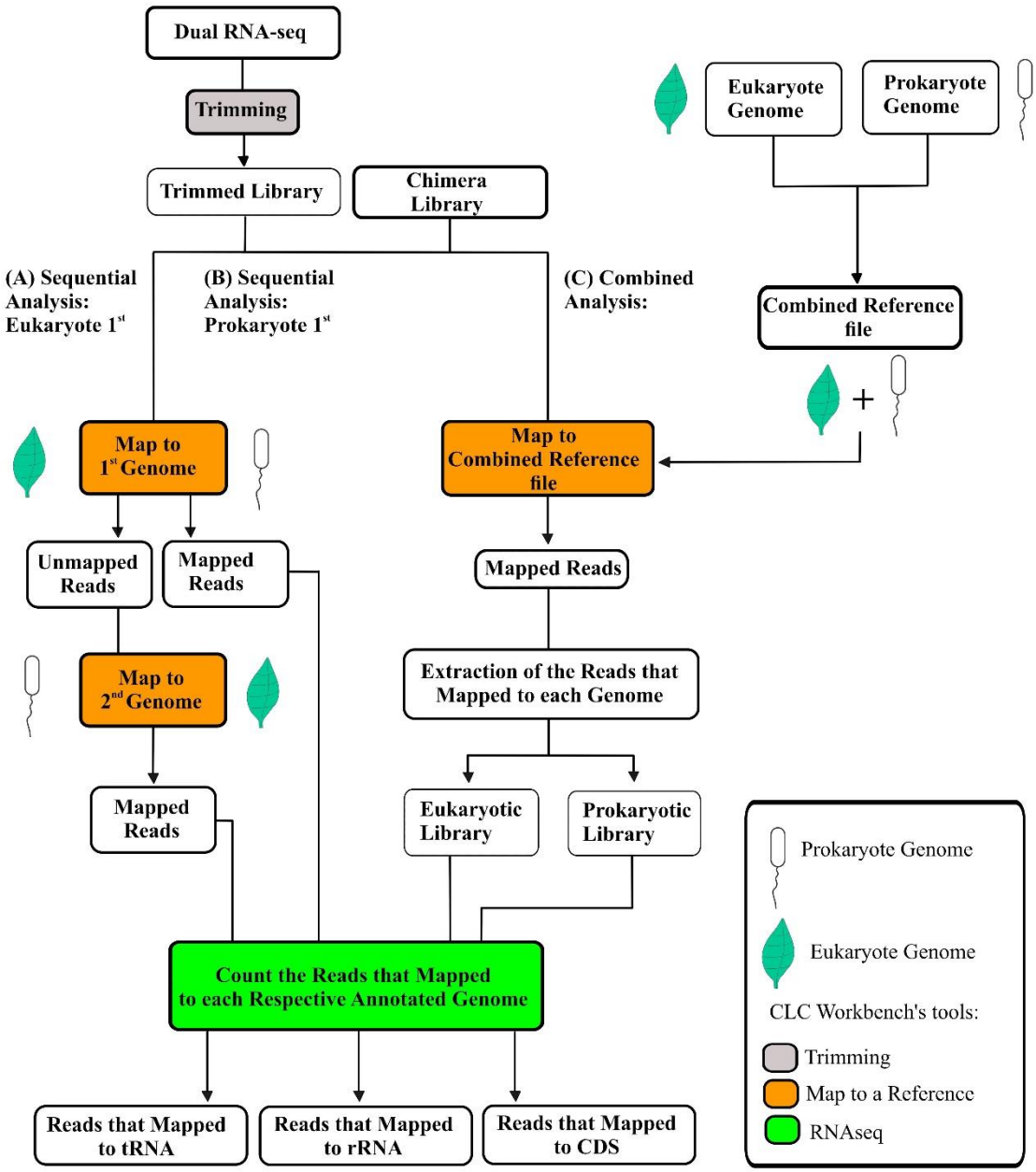
Supplementary material - the following online material is available for this article:

Figure S1. Filtering procedure and construction of the Chimera library. (A) Direct mapping used to filter the libraries to eliminate potential contamination reads (those that do not map to the respective genome). (B) Construction of the Chimera Library by merging the files with reads that mapped to the *Herbaspirillum seropedicae* and *Zea mays* reference genomes only.

640 **Figure S2.** Mapping strategies to determine the number of cross-mapping reads. (A)
641 Filtered libraries mapped to the reference genome of the other organism. (B) Filtered
642 libraries mapped to the Combined Genomes.
643
644 **Figure S3.** Step-by-step of how to extract the reads after the mapping step using the
645 Combined Reference in the CLC workbench environment. 1. Select the result file; 2.
646 Select the lines with the reads of the organism; 3. Click on "Extract Subset". In the
647 following window, select a place to save and the new file will contain all the reads that
648 mapped to that genome, without the need of any extra alignment step.
649
650 **Table S1.** Top 20 most counted loci whose reads were lost from the *Herbaspirillum* (A)
651 and *Z. mays* (B) libraries according to the mapping strategy used.
652
653 **Table S2.** Top 20 most counted loci whose reads were gained to the *Herbaspirillum* (A)
654 and *Z. mays* (B) libraries according to the mapping strategy used.
655
656 **Table S3**. Library features and number of total reads attributed to the *Bradyrhizobium*
657 *elkanii* or *Glycine max* genomes according to the mapping approach. BR16 and ER48:
658 soybean varieties BR16 and Embrapa 48, respectively.
659
660 **Table S4.** Library features and number of total reads attributed to the *Fusarium*
661 *verticillioides* or *Zea mays* genomes according to the mapping approach. Reads in Pairs
662 indicates the total number of reads that were aligned in pairs. Broken Pair Reads indicates
663 reads that could align independently; none of the possible placements of the pair satisfied

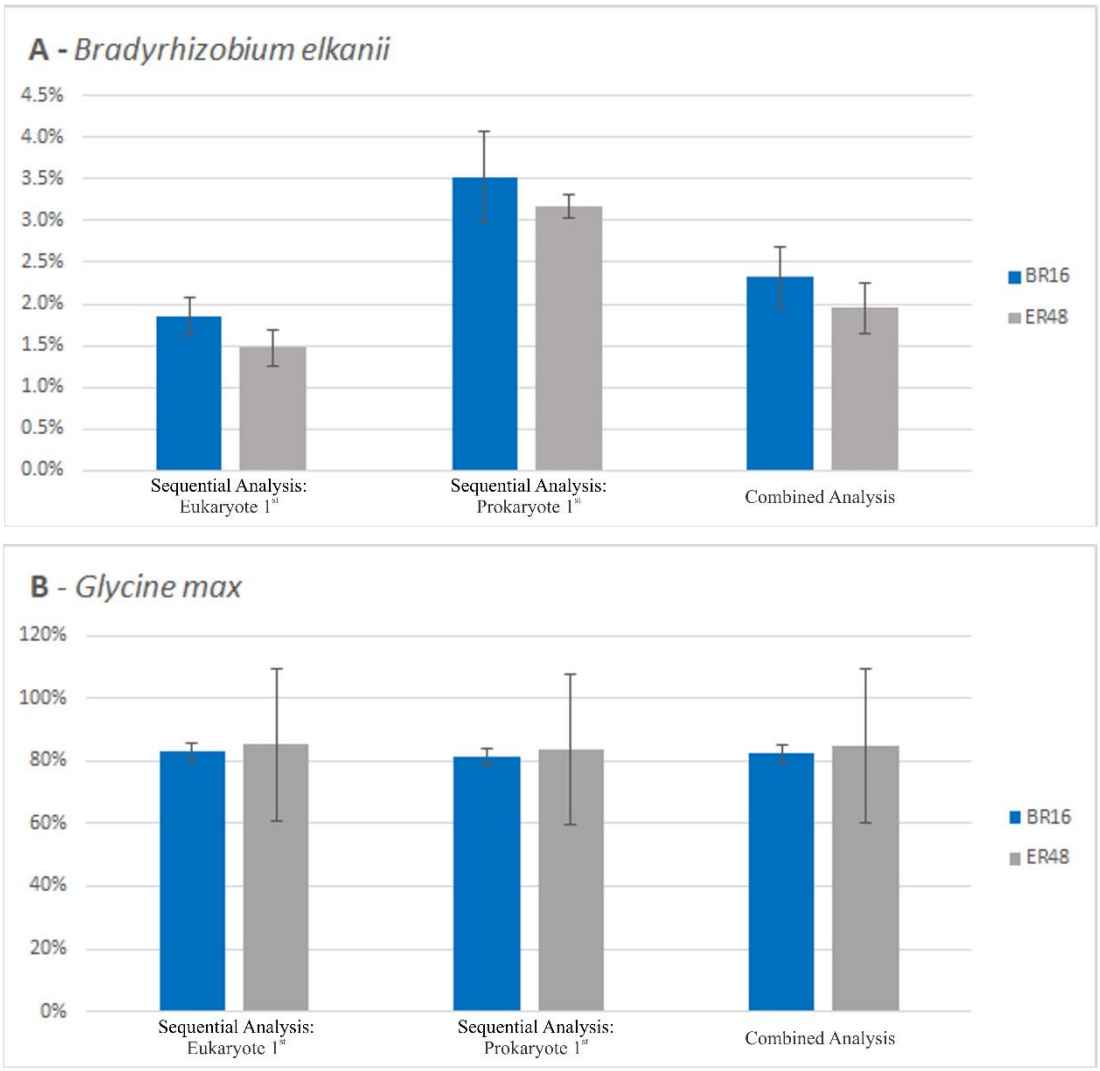
664 the pairing criteria. CO354: susceptible maize variety CO354 inoculated with *F.*
665 *verticillioides* from Lanubile *et al.* (2014).
666

667



668

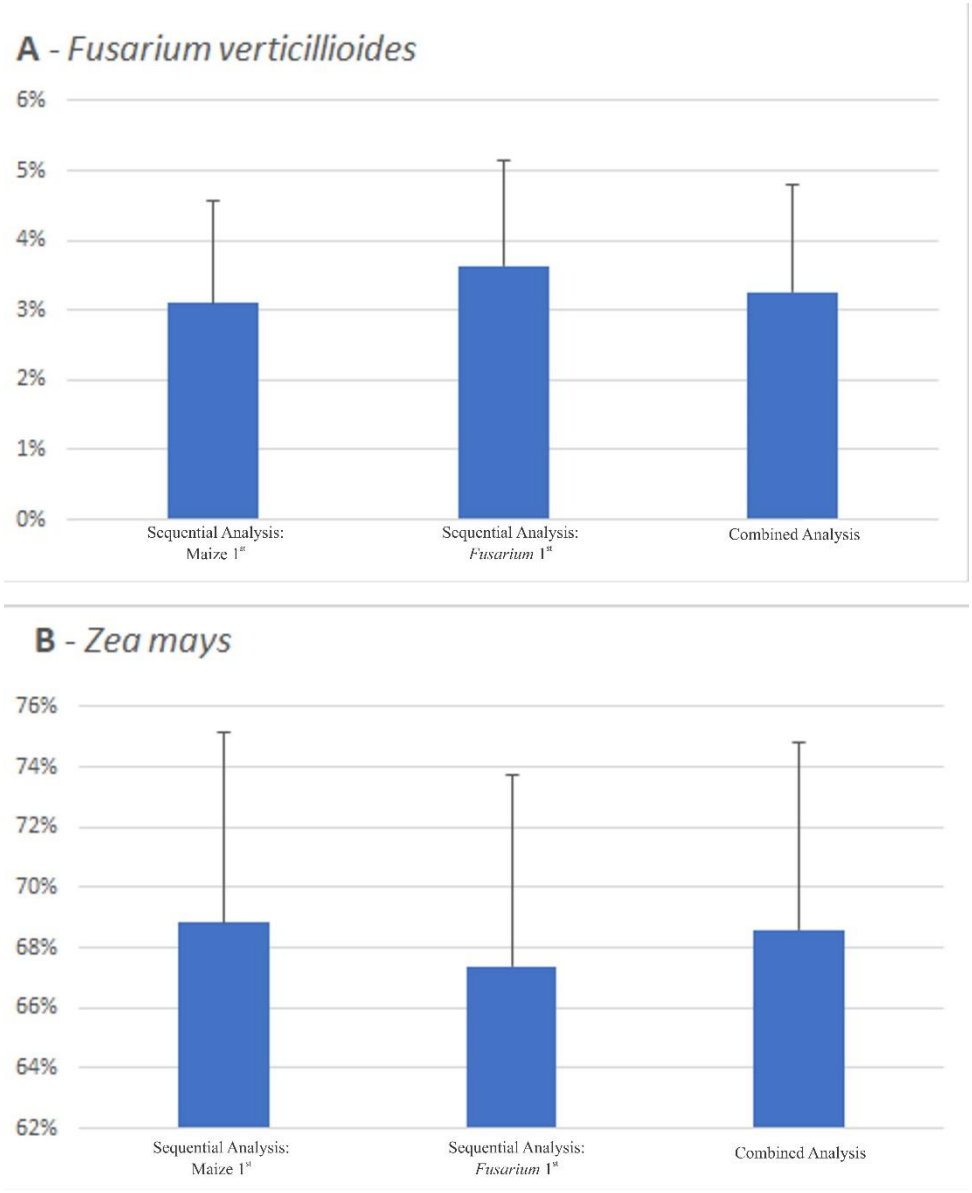
669 **Figure 1.** Mapping strategies for Dual RNA-Seq analysis. (A) Sequential analysis
670 aligning libraries to the eukaryotic genome first- Eukaryote 1st; (B) Sequential analysis
671 aligning libraries to the prokaryotic genome first- Prokaryote 1st; (C) Combined analysis.
672



673

674 **Figure 2.** Percentage of reads mapped to (A) *Bradyrhizobium elkanii* or (B) *Glycine max*
675 depending on the methodology used in the *Glycine max* – *Bradyrhizobium elkanii*
676 experiment. Bars indicate twice the Standard Error. BR16 and ER48: soybean varieties
677 BR16 and Embrapa 48, respectively.

678



679

680 **Figure 3.** Percentage of reads mapped to (A) *Fusarium verticillioides* or (B) *Zea mays*
681 depending on the methodology used in the *Zea mays –Fusarium verticillioides*
682 experiment. Bars indicate twice the Standard Error.

683

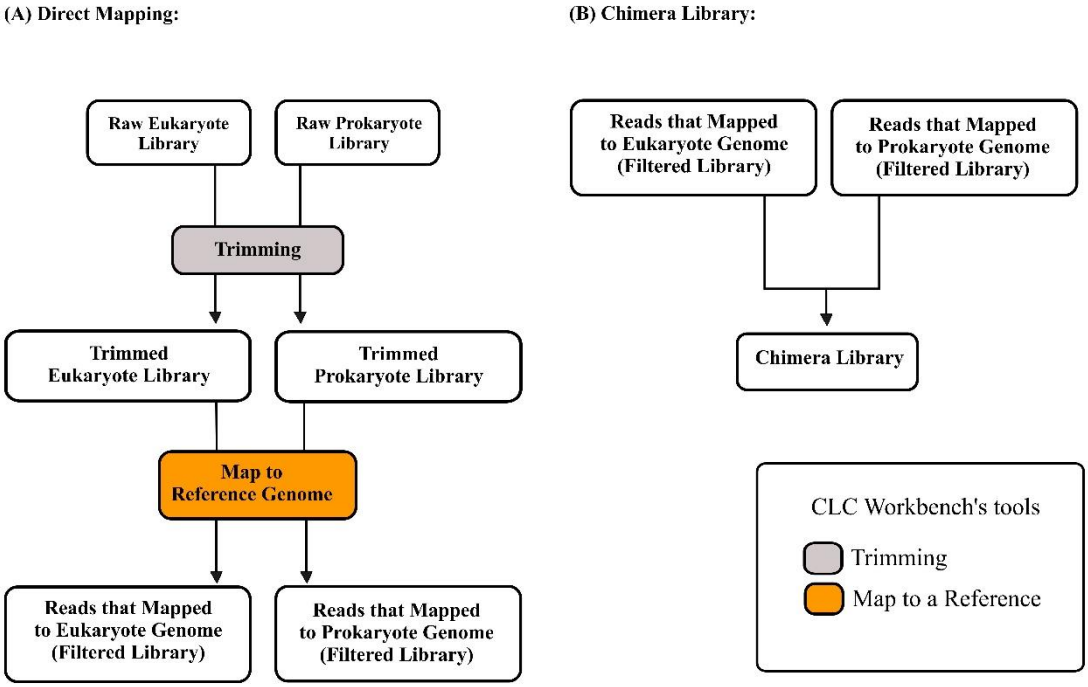
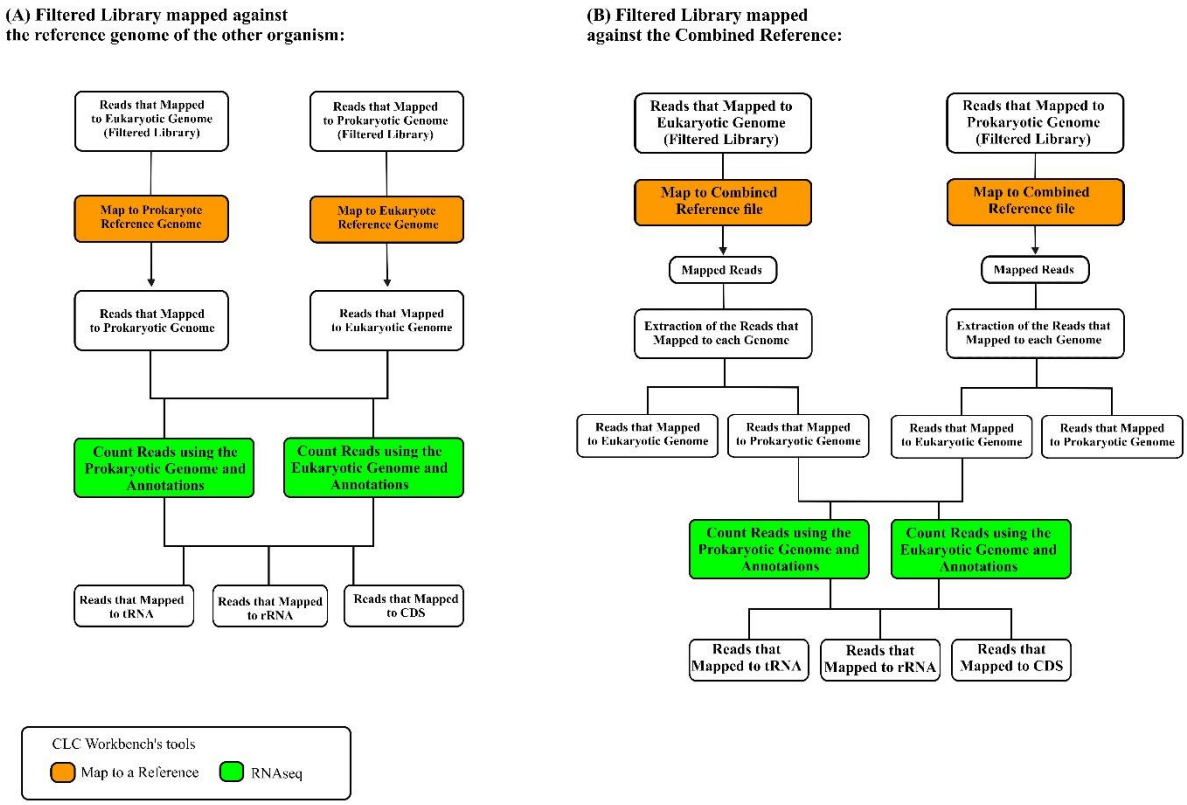


Figure S1. Filtering procedure and construction of the Chimera library. (A) Direct mapping used to filter the libraries to eliminate potential contamination reads (those that do not map to the respective genome). (B) Construction of the Chimera Library by merging the files with reads that mapped to the *Herbaspirillum seropedicae* and *Zea mays* reference genomes only.



691

692 **Figure S2.** Mapping strategies to determine the number of cross-mapping reads. (A)

693 Filtered libraries mapped to the reference genome of the other organism. (B) Filtered

694 libraries mapped to the Combined Genomes.

695

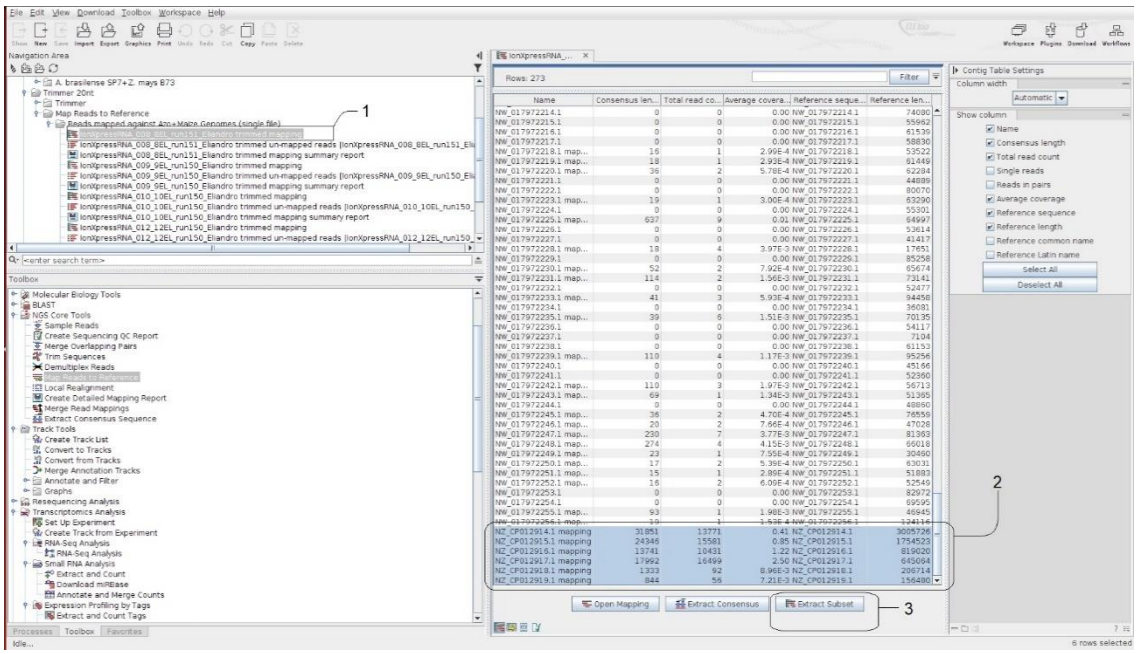


Figure S3. Step-by-step of how to extract the reads after the mapping step using the Combined Reference in the CLC workbench environment. 1. Select the result file; 2. Select the lines with the reads of the organism; 3. Click on “Extract Subset”. In the following window, select a place to save and the new file will contain all the reads that mapped to that genome, without the need of any extra alignment step.

Table 1. Library features and number of total reads attributed to the *Herbaspirillum seropedicae* or *Zea mays* genomes according to the mapping approach. The analyses were performed with the genomes without annotations. Cross-mapping: reads of each individual library were mapped to the reference genome of the other organism. Sequential Analysis: The library was first mapped to one reference genome, reads that fail to map to the first genome were mapped to the other genome. Eukaryote 1st/Prokaryote 1st indicates the first reference used. Combined analysis: libraries were mapped to a merged file containing both reference genomes (Combined Reference).

Library	Total Reads	Total reads after trimming	Number of Reads After Library filtration	Cross-Mapping		Mapping Strategy					
						Sequential Analysis				Combined Analysis	
						Eukaryote 1st		Prokaryote 1st			
				H. seropedicae	*Z. mays*	*H. seropedicae*	*Z. mays*	*H. seropedicae*	*Z. mays*	*H. seropedicae*	*Z. mays*
H. seropedicae	28,654,755	19,174,992	8,808,728	-	2,252,606	-	-	-	-	8,718,943	86,366
Z. mays	24,300,211	24,255,170	22,200,875	7,659	-	-	-	-	-	394	22,200,465
Chimera Library	-	-	31,009,603	-	-	6,556,122	24,453,481	8,816,387	22,193,216	8,719,342	22,286,826

715 **Table 2.** Number of reads mapped to tRNA, rRNA, and coding loci (CDS) according to the mapping methodology used.

716

Mapping Strategy	Library	Reference used to count the reads	Number of Reads Mapped to			Unmapped reads
			tRNA	rRNA	CDS loci	
Direct Mapping	*H. seropedicae* *Z. mays*	*H. seropedicae* *Z. mays*	235,428 1,692	5,649,357 3,003	2,307,178 21,051,646	616,765 1,144,534
Eukaryote 1st	Chimera Library	*H. seropedicae* *Z. mays*	193,088 13,175	4,263,732 403,522	1,616,947 22,119,380	482,355 1,917,404
Prokaryote 1st	Chimera Library	*H. seropedicae* *Z. mays*	235,430 1,686	5,650,140 2,255	2,313,367 21,045,092	617,450 1,144,183
Combined Analysis	Chimera Library	*H. seropedicae* *Z. mays*	234,656 1,739	5,618,104 11,372	2,258,318 21,104,107	608,264 1,169,608

717

718

719

720

721

Table 3. Comparison of the number of reads incorrectly mapped due to cross-mapping. Reads that incorrectly aligned to the reference genome were counted using the annotated genome indicated on the table. The unmapped reads are a result of the counting parameters that eliminate reads that aligned in more than five loci and in the intergenic regions. *CDS with at least 10 reads assigned to them. Exception made to the *Z. mays* library mapped to the Combined Reference, which refers to CDS with at least one read assigned to it.

Library	Reference used to map the reads	Reference used to count the cross-mapped reads	Number of reads mapped to			CDS*	Unmapped reads
			tRNA	rRNA	CDS Loci		
H. seropedicae	*Z. mays*	*H. seropedicae*	42,340	1,385,625	690,231	3,690	134,410
		Z. mays	11,483	400,519	1,067,734	10,715	772,870
	Combined Reference	*H. seropedicae*	748	30,608	46,880	880	8,130
		Z. mays	47	8,685	52,530	575	25,104
Z. mays	*H. seropedicae*	*Z. mays*	6	748	6,554	72	351
		H. seropedicae	2	783	6,189	65	685
	Combined Reference	*Z. mays*	0	308	57	43	29
		H. seropedicae	0	329	59	49	6

729 **Table 4**. Number of reads mapped to tRNA, rRNA, and coding loci (CDS) according to
730 the mapping methodology and experiment used. BR16 and ER48: soybean varieties BR16
731 and Embrapa 48, respectively. CO354: susceptible maize variety CO354 inoculated with
732 *F. verticillioides* from Lanubile *et al.* (2014).
733

Samples	Biological Replicate	Mapping Strategy	Reference used to count the reads	Number of Reads Mapped to			Unmapped reads
				tRNA	rRNA	CDS loci	
Soybean + Bradyrhizobium:							
BR16	I	Eukaryote 1st	*G. max*	9,262	458,414	6,891,692	275,078
	II		*G. max*	14,526	524,794	7,071,502	298,956
	I		*B. elkanii*	6,140	137,368	18,831	1,349
	II		*B. elkanii*	8,431	153,677	21,122	2,025
ER48	I		*G. max*	7,486	284,116	4,272,465	161,811
	II		*G. max*	12,566	400,613	5,612,058	263,977
	I		*B. elkanii*	5,423	87,283	8,411	1,057
	II		*B. elkanii*	5,219	64,293	17,133	1,401
BR16	I	Prokaryote 1st	*G. max*	7,784	337,378	6,875,036	271,697
	II		*G. max*	11,509	383,520	7,050,057	293,615
	I		*B. elkanii*	8,597	262,024	31,914	3,704
	II		*B. elkanii*	13,152	301,072	35,907	6,201
ER48	I		*G. max*	5,803	209,395	4,254,381	158,528
	II		*G. max*	10,145	321,792	5,577,283	258,616
	I		*B. elkanii*	8,443	165,541	21,574	4,387
	II		*B. elkanii*	9,329	147,323	45,409	7,363
BR16	I	Combined Analysis	*G. max*	8,571	420,015	6,889,137	273,168
	II		*G. max*	13,227	477,056	7,067,921	295,728
	I		*B. elkanii*	7,097	172,756	19,253	1,595
	II		*B. elkanii*	10,299	200,206	21,842	2,236
ER48	I		*G. max*	6,617	252,998	4,269,474	159,801
	II		*G.max*	11,597	378,032	5,608,032	260,854
	I		*B. elkanii*	6,907	118,911	9,057	1,166
	II		*B. elkanii*	6,872	89,370	18,201	1,613

(Table 4-continuation) *Maize + Fusarium:*

	I		*Z. mays*	257	46,053	68,625,513	1,973,035
	II		*Z. mays*	280	43,537	69,135,019	2,015,545
	III		*Z. mays*	279	36,277	59,728,346	1,690,684
CO354	I	Maize 1st	*F. verticillioides*	47	509	2,838,639	150,957
	II		*F. verticillioides*	23	293	1,543,960	83,640
	III		*F. verticillioides*	53	526	3,664,696	196,770
	I		*Z. mays*	257	35,577	68,045,490	1,678,420
	II		*Z. mays*	280	35,190	68,601,878	1,713,630
	III		*Z. mays*	279	26,101	59,171,896	1,446,308
CO354	I	*Fusarium* 1st	*F. verticillioides*	48	583	3,415,917	458,718
	II		*F. verticillioides*	23	369	2,074,133	396,794
	III		*F. verticillioides*	53	601	4,218,730	453,663
	I		*Z. mays*	257	38,312	68,526,796	1,960,084
	II		*Z. mays*	280	37,882	69,060,416	2,006,113
	III		*Z. mays*	279	28,336	59,626,602	1,677,822
CO354	I	Combined Analysis	*F. verticillioides*	48	516	2,850,971	166,912
	II		*F. verticillioides*	23	291	1,538,950	94,107
	III		*F. verticillioides*	53	533	3,690,721	213,560

734

735

736

737

738 **Table S1.** Top 20 most counted loci whose reads were lost from the *Herbaspirillum* (A) and *Z. mays* (B) libraries according to the mapping
739 strategy used.
740 (A)

Mapping Strategy	Locus tag	Annotation	RPKM	Unique gene reads	Total gene reads
Cross-Mapping	ssrA_2	miscRNA (Gene Ids: 31910439 and 31910466)	139,688	36,690	36,735
	ACP92_RS13610	hypothetical protein	77,769	18,327	18,358
	ACP92_RS21415	porin	19,808	15,771	15,914
	ACP92_RS17420	NADPH: quinone reductase	19,376	13,733	13,762
	ACP92_RS01180	acetyl-CoA acetyltransferase	12,352	9,989	10,026
	ACP92_RS14915	chemotaxis protein CheW	28,367	9,927	9,927
	ACP92_RS18115	nitrate reductase subunit alpha	3,811	9,717	9,810
	ACP92_RS19205	amino acid ABC transporter substrate-binding protein	8,917	5,804	5,835
	ACP92_RS18120	nitrate reductase subunit beta	4,810	5,140	5,159
	ACP92_RS14850	2-oxoglutarate dehydrogenase subunit E1	2,554	5,040	5,051
	ACP92_RS18395	membrane protein	11,467	4,841	4,844
	ACP92_RS12875	endopeptidase La	2,828	4,691	4,702
	ACP92_RS06995	ATP-dependent Clp protease ATP-binding subunit ClpA	2,688	4,267	4,274
	ACP92_RS19810	ribosomal subunit interface protein	16,907	4,229	4,236
	ACP92_RS23345	hypothetical protein	12,796	4,211	4,213
	ACP92_RS14905	aconitate hydratase	2,228	4,148	4,170
	ACP92_RS23935	ABC transporter permease	4,877	4,034	4,070
	ACP92_RS06455	filamentous hemagglutinin protein	736	4,029	4,052
	ACP92_RS17300	branched chain amino acid ABC transporter substrate-binding protein	4,062	3,200	3,205
	ACP92_RS23135	succinyl-CoA--3-ketoacid-CoA transferase	6,788	3,046	3,050

Combined	ACP92_RS13610	hypothetical protein	415,581	6,654	6,663
	ACP92_RS15580	alkene reductase	10,753	543	552
	ACP92_RS14850	2-oxoglutarate dehydrogenase subunit E1	3,082	407	414
	ACP92_RS01180	acetyl-CoA acetyltransferase	7,038	368	388
	ACP92_RS17425	hypothetical protein	23,018	326	382
	ACP92_RS17420	NADPH: quinone reductase	7,794	359	376
	ssrA_2	miscRNA (Gene Ids: 31910439 and 31910466)	17,972	286	321
	ACP92_RS18120	nitrate reductase subunit beta	4,228	296	308
	ACP92_RS18115	nitrate reductase subunit alpha	1,590	248	278
	ACP92_RS11335	hypothetical protein	3,884	239	242
	ACP92_RS06190	protein translocase subunit SecD	2,699	231	241
	ACP92_RS12875	endopeptidase La	2,099	233	237
	ACP92_RS21415	porin	3,812	197	208
	ACP92_RS12540	isocitrate dehydrogenase, NADP-dependent	1,828	188	191
	ACP92_RS05975	hypothetical protein	10,722	189	190
	ACP92_RS20135	hypothetical protein	6,580	186	186
	ACP92_RS23935	ABC transporter permease	3,264	176	185
	ACP92_RS05425	hypothetical protein	158	102	176
	ACP92_RS18550	transcriptional regulator	5,903	175	176
	ACP92_RS11435	hypothetical protein	12,697	175	175

745 (B)

Mapping Strategy	Locus Tag	Description	RPKM	Unique gene reads	Total gene reads
Cross-Mapping	LOC103629174	trihelix transcription factor GT-3b	8,504	114	279
	LOC542551	60S acidic ribosomal protein P2A-like	20,418	269	277
	LOC542323	acidic ribosomal protein P2a (ARPP2A)	18,561	260	263
	pco106379	RNA-binding (RRM/RBD/RNP motifs) family protein	15,534	182	208
	LOC542725	uncharacterized LOC542725	33,925	192	203
	LOC542337	acidic ribosomal protein P2a-2 (ARPP2A-2)	12,514	165	166
	LOC103627451	ribosomal RNA-processing protein 14	15,394	125	147
	fdh1	formaldehyde dehydrogenase homolog 1 (fdh1)	5,476	131	131
	LOC100194217	responsive to abscisic acid15	21,139	115	124
	pco102485	uncharacterized LOC100282592	2,842	90	93
	LOC103638017	ethylene-responsive transcription factor RAP2-1	9,638	60	77
	LOC100286241	heat shock 70 kDa protein	3,667	71	71
	LOC100194313	acidic ribosomal protein P2a-4	5,246	68	71
	HMG13	high mobility group protein 1	4,573	63	63
	LOC100193124	myosin-like protein	1,537	58	58
	gpm583	Guanine nucleotide-binding protein beta subunit-like protein	3,351	49	49
	LOC103639924	probable mediator of RNA polymerase II transcription subunit 37c	1,852	44	47
	LOC100501446	uncharacterized LOC100501446	12,593	0	43
	LOC109939197	NADP-dependent malic enzyme, chloroplastic	1,132	0	41
	LOC103650526	probable mediator of RNA polymerase II transcription subunit 37c	1,534	36	40
Combined	pco106379	RNA-binding (RRM/RBD/RNP motifs) family protein	42,937	1	5

	LOC103648227	AAA-ATPase ASD mitochondrial	15,256	1	2
	LOC109944797	cold and drought-regulated protein CORA	34,983	0	2
	LOC109940833	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic-like	3,499	0	2
	LOC103632470	aminotransferase ALD1 homolog	3,736	1	1
	LOC103634426	probable protein phosphatase 2C 31	5,291	1	1
	LOC100502427	uncharacterized LOC100502427	1,606	1	1
	umc2536	uncharacterized LOC100275320	20,689	0	1
	LOC100281029	GIR1	22,873	1	1
	LOC103650526	probable mediator of RNA polymerase II transcription subunit 37c	4,410	1	1
	LOC100272685	uncharacterized LOC100272685	6,075	1	1
	LOC100194009	uncharacterized LOC100194009	1,085	1	1
	LOC100191812	uncharacterized LOC100191812	1,003	1	1
	LOC103651416	DUF538 family protein	24,265	0	1
	LOC103655071	probable pectinesterase/pectinesterase inhibitor 46	34,000	0	1
	LOC100275424	uncharacterized LOC100275424	7,843	0	1
	LOC100275783	uncharacterized LOC100275783	5,362	0	1
	LOC542333	cytosolic glyceroldehyde-3-phosphate dehydrogenase (GAPC3)	5,438	0	1
	LOC100191329	uncharacterized LOC100191329	4,888	0	1
	LOC100279325	meiotic recombination protein SPO11	2,499	1	1

Unique gene reads = reads that mapped on a single locus; Total gene reads= sum of unique gene reads and reads that mapped to more than 5 loci; RPKM = reads per kilobase million.

751 **Table S2.** Top 20 most counted loci whose reads were gained to the *Herbaspirillum* (A) and *Z. mays* (B) libraries according to the mapping
752 strategy used.
753 (A)

Mapping Strategy	Locus Tag	Description	RPKM	Unique gene reads	Total gene reads
Cross-Mapping	ACP92_RS15225	cell wall surface anchor protein	2,544	333	339
	ACP92_RS16690	ABC transporter substrate-binding protein	36,149	228	298
	ACP92_RS07965	heavy metal translocating P-type ATPase	19,151	232	277
	ACP92_RS19745	single-stranded DNA-binding protein	42,097	133	136
	ACP92_RS01655	UDP-N-acetylmuramoyl-L-alanine--D-glutamate ligase	9,772	116	129
	ACP92_RS13570	S-(hydroxymethyl)glutathione dehydrogenase	18,829	129	129
	ACP92_RS04965	molecular chaperone GroEL	12,339	126	126
	ACP92_RS23590	Ni/Fe hydrogenase	30,098	96	114
	ACP92_RS01565	MexE family multidrug efflux RND transporter periplasmic adaptor subunit	13,195	93	98
	ACP92_RS10470	phosphatidylinositol kinase	11,649	72	93
	ACP92_RS18805	two-component sensor histidine kinase	4,412	77	77
	ACP92_RS05425	hypothetical protein	477	68	70
	ACP92_RS16420	colicin V production protein	21,931	65	68
	ACP92_RS00385	serine/threonine protein phosphatase	5,105	60	60
	ACP92_RS07325	hypothetical protein	12,673	60	60
	ACP92_RS03935	EscV/YscV/HrcV family type III secretion system export apparatus protein	3,792	42	49
	ACP92_RS12400	glyoxylate/hydroxypyruvate reductase A	7,911	26	47
	ACP92_RS17960	23S rRNA pseudouridylate synthase B	3,506	44	44
	ACP92_RS17105	RNA helicase	4,216	37	40
	ACP92_RS13260	membrane protein	3,993	39	39

Combined	ACP92_RS16605	hypothetical protein	313,873	5	5
	ACP92_RS19745	single-stranded DNA-binding protein	129,879	4	4
	ACP92_RS05425	hypothetical protein	1,428	2	2
	ACP92_RS07830	FAD-dependent oxidoreductase	22,735	2	2
	ACP92_RS20885	hypothetical protein	61,746	2	2
	ACP92_RS00010	DNA polymerase III subunit beta	15,311	1	1
	ACP92_RS00220	hypothetical protein	41,850	1	1
	ACP92_RS00520	DNA-directed RNA polymerase subunit beta	4,127	1	1
	ACP92_RS00925	hypothetical protein	10,992	1	1
	ACP92_RS24015	hypothetical protein	19,152	1	1
	ACP92_RS03200	IMP dehydrogenase	5,004	1	1
	ACP92_RS03310	hypothetical protein	27,831	1	1
	ACP92_RS03695	hypothetical protein	26,903	1	1
	ACP92_RS04410	phage late control protein	13,780	1	1
	ACP92_RS04995	transcriptional regulator	18,225	1	1
	ACP92_RS05310	Ribulo kinase	10,291	1	1
	ACP92_RS05790	IclR family transcriptional regulator	21,400	1	1
	ACP92_RS05930	branched-chain amino acid ABC transporter permease	18,225	0	1
	ACP92_RS06930	DNA-binding response regulator	24,458	1	1
	ACP92_RS07090	type I glyceraldehyde-3-phosphate dehydrogenase	16,715	1	1

758 (B)

Mapping Strategy	Locus Tag	Description	RPKM	Unique gene reads	Total gene reads
Cross-Mapping	LOC103641534	Protein RAE1	4,942	98,651	112,018
	LOC542347	thylakoid assembly 1	3,610	70,795	84,977
	LOC103630900	uncharacterized LOC103630900	1,253	30,530	37,387
	LOC100272896	hypothetical protein	2,429	15,604	15,737
	LOC103626304	protein ACTIVITY OF BC1 COMPLEX KINASE 7, chloroplastic	262	1,729	15,401
	LOC103649599	cationic amino acid transporter 5	6,773	15,194	15,274
	LOC100284199	nonspecific lipid-transfer protein	15,081	11,790	11,819
	cl9362_1	MAK16-like protein RBM13	1,855	9,317	9,869
	LOC100277826	Tubulin binding cofactor C domain-containing protein	513	8,195	9,744
	LOC100382570	uncharacterized LOC100382570	769	6,586	8,587
	LOC100304272	Vesicle-associated protein 2-1	2,581	8,199	8,207
	LOC100285109	uncharacterized LOC100285109	943	6,598	7,659
	LOC100277664	uncharacterized LOC100277664	2,904	7,110	7,500
	LOC103626482	callose synthase 3	95	5,564	7,240
	LOC103632082	tRNA ligase 1	213	124	6,854
	LOC103643067	endoplasmic reticulum-Golgi intermediate compartment protein 3	912	5,924	6,529
	LOC100191302	uncharacterized LOC100191302	428	5,261	5,339
	pco103153	Nicotinate phosphoribosyltransferase 2	337	4,901	5,140
	LOC109940494	uncharacterized LOC109940494	1,448	49	5,086
	LOC100382629	uncharacterized LOC100382629	277	4,697	4,897
Combined	LOC103649599	cationic amino acid transporter 5	45,185	4,981	5,013

	LOC103641534	Protein RAE1	3,187	752	3,554
	LOC100274325	Putative polyol transporter 1	363	553	556
	LOC103626482	callose synthase 3	99	355	372
	cl9072_1(438)		1,410	184	327
	LOC100384061	uncharacterized LOC100384061	3,434	281	285
	LOC100279342	E3 ubiquitin-protein ligase COP1	737	77	271
	LOC100272560	transferase	2,030	214	238
	LOC542347	thylakoid assembly 1	177	147	205
	LOC103634584	ncRNA	2,332	197	197
	gpm679	uncharacterized LOC100280372	253	6	197
	LOC100276350	CASP-like protein 13	107	186	189
	LOC100276421	GW2	635	12	184
	LOC100278614	hypothetical protein	2,225	180	180
	LOC103648897	tRNA ligase 1-like	31	158	177
	LOC100285411	embryogenesis transmembrane protein	2,406	0	173
	aox1_1	alternative oxidase 1 (GeneID: 542073 and 542074)	1,132	2	159
	LOC100384715	hypothetical protein	1,651	158	159
	LOC100283538	cyclin delta-2	1,127	99	156
	LOC103650887	PTI1-like tyrosine-protein kinase 3	616	152	154

Unique gene reads = reads that mapped on a single locus; Total gene reads= sum of unique gene reads to reads that mapped to more than 5 loci;
RPKM = reads per kilobase million.

764 **Table S3**. Library features and number of total reads attributed to the *Bradyrhizobium elkanii* or *Glycine max* genomes according to the mapping
765 approach. BR16 and ER48: soybean varieties BR16 and Embrapa 48, respectively.
766

Samples	**Biological Replicate**	**Total Reads**	**Total reads after trimming**	Mapping Strategy								
				Sequential Analysis						Combined Analysis		
				Eukaryote 1st			Prokaryote 1st					
				B. elkanii	*G. max*	Unmapped	*B. elkanii*	*G. max*	Unmapped	*B. elkanii*	*G. max*	Unmapped
BR16	I	9,065,000	9,030,791	163,688	7,634,446	1,232,657	306,239	7,491,895	1,232,657	200,701	7,590,891	1,239,199
	II	9,792,000	9,750,645	185,255	7,909,778	1,655,612	356,332	7,738,701	1,655,612	234,583	7,853,932	1,662,130
ER48	I	5,693,000	5,663,651	102,174	4,725,878	835,599	199,945	4,628,107	835,599	136,041	4,688,890	838,720
	II	7,308,000	7,255,860	88,046	6,289,214	878,600	209,424	6,167,836	878,600	116,056	6,258,515	881,289

767
768
769
770
771
772
773
774
775

Table S4. Library features and number of total reads attributed to the *Fusarium verticillioides* or *Zea mays* genomes according to the mapping approach. Reads in Pairs indicates the total number of reads that were aligned in pairs. Broken Pair Reads indicates pair-reads that could be aligned independently, none of the possible placements of the pair satisfies the pairing criteria. CO354: susceptible maize variety CO354 inoculated with *F. verticillioides* from Lanubile *et al.* (2014).

Sample	**Biological Replicate**	**Total reads after trimming**	Sequential Analysis							
			Maize 1st				*Fusarium* 1st			
			F. verticillioides		*Z. mays*		*F. verticillioides*		*Z. mays*	
			Reads in pairs	Broken paired reads	Reads in pairs	Broken paired reads	Reads in pairs	Broken paired reads	Reads in pairs	Broken paired reads
CO354	I	83,693,936	2,638,310	351,842	57,503,900	13,140,958	3,073,034	802,232	55,856,108	13,903,636
	II	82,742,688	1,412,600	215,316	58,018,836	13,175,545	1,790,488	680,831	57,277,376	13,073,602
	III	74,252,520	3,422,402	439,643	50,135,992	11,319,594	3,887,094	785,953	49,000,248	11,644,336

Sample	Biological Replicate	Total reads after trimming	Combined Analysis				
			F. verticillioides		*Z. mays*		
			Reads in pairs	Broken paired reads	Reads in pairs	Broken paired reads	Total Reads
CO354	I	83,693,936	2,756,392	262,055	57,223,066	13,302,383	73,543,896
	II	82,742,688	1,468,926	164,445	57,841,826	13,262,865	72,738,062
	III	74,252,520	3,604,734	300,133	50,026,720	11,306,319	65,237,906

5. Artigo 2: The genetic interaction between *Azospirillum brasilense* and maize during the inhibition of indole-3-acetic acid production by the plant

The genetic interaction between *Azospirillum brasilense* and maize during the inhibition of indole-3-acetic acid production by the plant

Eliandro Espindula[1], Edilena Reis Sperb[1], Vânia Carla Silva Pankievicz[2], Thalita Regina Tuleski[2], Michelle Zibetti Tadra-Sfeir[2], Paloma Bonato[2], Emanuel Maltempi de Souza[2], Luciane Maria Pereira Passaglia[1]*

[1] Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

[2] Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná (UFPR), Centro Politécnico, Curitiba, PR, Brazil

* Corresponding author: LMP Passaglia, Departamento de Genética, IB-UFRGS, Av. Bento Gonçalves 9500, 91501-970 Porto Alegre, Brazil. Phone: +5551 3308 9813; e-mail: luciane.passaglia@ufrgs.br

Artigo a ser submetido à revista Molecular Plant-Microbe Interaction

Abstract

Maize is one of the most used plant species in human and animal food. For this reason, ways to increase its production without increasing the planted area are always necessary. One way of achieving this is by using plant growth-promoting bacteria (PGPB). Among the most well-known PGPB are the members of the genus *Azospirillum*, which have as main characteristics the ability to fix nitrogen and to produce phytohormones. This work aimed to better understand the genetic relationship between *Azospirillum brasilense* and maize during the inhibition of indole-3-acetic acid (IAA) production by the plants. For this, the chemical compound Yucasin [5-(4-chlorophenyl)-4H-1,2,4-triazole-3-thiol], an inhibitor of one of the IAA plant-production pathways, was applied. Dual RNA-Seq was used to analyze the gene expression of the bacterium and maize using a combined analysis approach. With this methodology was possible to observe maize genes involved in the ABA response pathway being differentially expressed. Their pattern of expression suggested that the plant was responding to the presence of the Yucasin. Coding genes for the transcription factors MYB29 and WRKY71, involved in biotic and abiotic stress responses, were up-regulated in some experimental conditions, indicating that the plant was responding to the presence of the bacterium and the Yucasin. Concerning the bacterial gene expression, genes involved in the transmembrane transport and terpenoid backbone biosynthesis were differentially expressed.

Keywords: Dual RNA-Seq; *Azospirillum brasilense*; *Zea mays*; Plant Growth-Promoting Bacteria; IAA.

Introduction

Plant Growth-Promoting Bacteria (PGPB) are a group of beneficial microorganisms that can colonize the rhizosphere, the phyllosphere, the root's surface and plant's internal tissues (Hungria *et al.* 2010) stimulating plant growth (Verna *et al.* 2010). It is believed that these bacteria can promote plant growth through the combination of several abilities, like biological nitrogen fixation, production of phytohormones (especially indole-3-acetic acid – IAA), vitamins and growth factors (Babalola, 2010; Hungria *et al.* 2010; Bashan and de-Bashan, 2010).

Azospirillum are gram-negative, free-living bacteria that can be isolated from the rhizosphere of grasses and cereals grown in both tropical and temperate climates (Steenhoudt and Vanderleyden, 2000). Under micro anaerobic conditions and in the absence of available nitrogen, these bacteria can fix the atmospheric nitrogen and convert it into ammonia. This ammonia is secreted into the environment and readily absorbed by the plants (Cassán *et al.* 2009; Steenhoudt and Vanderleyden, 2000; Fornasieri, 1992). Besides fixing atmospheric nitrogen, bacteria of the genus *Azospirillum* also produce phytohormones (IAA, gibberellins, ethylene, and polyamines). It is believed that these two characteristics are responsible for stimulating the increase of the final dry mass of the plants (Perrig *et al.* 2007; Bashan and de-Bashan, 2010).

Among the phytohormones produced by PGPB and plants, auxins (mainly IAA) are the most studied (Spaepen *et al.* 2007; Baudoin, 2010; Spaepen and Vanderleyden, 2010; Yue *et al.* 2014). In both plant and PGPB, the main biosynthesis pathway uses tryptophan (Trp) as a precursor for IAA synthesis (Spaepen *et al.* 2007). In bacteria, IAA is used as a signaling molecule that allows them to stimulate root growth and carbohydrate exudation by plants helping to create a favorable environment for bacterial growth. Besides, this molecule is involved in the quorum-sensing process of bacteria, which permits them to control their activities based on population density (Spaepen *et al.* 2007; Duca *et al.* 2014; Yue *et al.* 2014). In plants, the auxins are responsible for regulating various aspects of their development, such as cell growth and differentiation, the establishment of apical dominance, differentiation of xylem, suppression of abscission, and formation of apical and root meristem (Bishopp *et al.* 2006; Yue *et al.* 2014).

Over the years, several IAA biosynthesis pathways from Trp have been proposed (Zhao, 2010). In these studies, enzymes belonging to the Tryptophan Amino-transferase of *Arabidopsis* (TAA) and the YUC-flavin monooxygenases families were positioned on different metabolic pathways (Zhao, 2010). However, recent studies have indicated that these two enzymes are part of the main pathway of IAA production in plants: TAA converts tryptophan to indole – 3 – pyruvate

(IPA), which is converted into IAA by the YUC (Mashiguchi *et al.* 2011; Won *et al.* 2011; Zhao, 2012; Zhao, 2014; Yue *et al.* 2014). It was observed by Nishimura *et al.* (2014) that the substance known as Yucasin [5-(4-chlorophenyl)-4H-1,2,4-triazol-3-thiol] is a competitive inhibitor of IPA, preventing its decarboxylation by YUC. Since Yucasin synergistically inhibits the production of IAA by this route, it can be used in studies of modulation of IAA production by the plant over time (Zhao, 2014). Finally, even though both bacteria and plants use Trp as the precursor to IAA biosynthesis, the enzymes involved in the IAA biosynthesis pathway in plants that are homologs to those of PGPB are still not know. This fact indicates that the IAA production pathway has evolved independently in bacteria and terrestrial plants (Yue *et al.* 2014). This convergent evolution points out the importance of this substance in both groups.

In the present work, the interaction between *Azospirillum brasilense* FP2 and maize during the inhibition of the maize TAA/YUC pathway by Yucasin was investigated. To achieve this objective, the genetic expression of the bacterium and plant was analyzed by Dual RNA-Seq. Data obtained from the sequencing of both plant and bacterium transcriptomes were analyzed using a combined analysis approach. With this methodology and after the counting step with the respective annotated genomes, it was possible to identify genes differentially expressed in both organisms.

Results

Bacterial growth in the presence of Yucasin, physiological data and scanning electron microscopy

To verify if the IAA inhibitor synthesis Yucasin could interfere with the growth and IAA production of *A. brasilense* FP2, bacterial growth curves were conducted in King B medium supplemented with this substance. The results showed that the presence of Yucasin in the growth medium affected neither the IAA production by *A. brasilense* FP2 (Figure 1) nor the bacterial growth (Figure S1).

To evaluate the physiological effects of Yucasin over the plant development and if the presence of *Azospirillum brasilense* FP2 could mitigate these effects, maize seedlings were subjected to Yucasin and *Azospirillum* treatments. After 10 and 15 days after inoculation, the lengths of aerial part and roots, and the number of lateral roots were measured (Figure 2). Concerning the root lengths, plantlets from group IV, which consisted of both Yucasin and *Azospirillum* treatments, presented higher lengths than those from groups I and II (with Yucasin treatment) (Figure 2A). This result indicated that the presence of *Azospirillum* in the group IV was

capable to stimuli the root growth even in the presence of the Yucasin, and that Yucasin did not interfere with the bacterial growth-promoting effect.

Differences in the number of lateral roots five days after plantlet's treatment with Yucasin were also observed (Figure 3). According to these results, Yucasin effectively inhibited the formation of lateral roots in plantlets of group II (with Yucasin). This group presented a lower number of lateral roots than plants from the group I (Figure 3). *A. brasilense* presence was enough to recovery the phenotype of plantlets of group IV (with Yucasin and bacterium). When the numbers of lateral roots of plantlets of groups II and IV were compared, the last presented a higher number than the first. Both results can also be observed in Figure S2. Plantlets from group II visually showed fewer lateral roots than plantlets from groups I and IV.

To confirm the presence of *A. brasilense* FP2 on maize's roots surface, a scanning electron microscopy analysis was performed using samples from biological replicates of the groups III and IV from the first sampling time point (10 DAI). It could be observed the presence of the bacterium on the surface of both experimental groups (Figure S3). One colony of *A. brasilense* inside the roots of plantlets of group IV, in the apoplastic region, was also observed. This indicated that the bacterium could effectively colonize the interior of maize roots (Figure S3C).

Transcriptome analysis

Since the presence of *A. brasilense* stimulated the formation of lateral roots in the plantlets in the presence of Yucasin, the gene expression pattern in all experimental groups were analyzed. Total RNA was extracted from each biological replicate of each experimental group, generating 12 RNA samples. Table 1 shows the main results obtained for each experimental group after the first bioinformatics analyzes.

Reads that mapped to the *A. brasilense* and *Z. mays* reference genomes were extracted from libraries of groups III and IV, and, at the counting step, those libraries were aligned against their respective reference genome with annotations. All reads that aligned in the intergenic regions, tRNA, and rRNA sequences were eliminated from the libraries, and the reads that aligned to coding sequences (CDS) were evaluated. As can be observed in Table 2, the majority of reads that mapped to the maize genome corresponded to multi-reads. On the other hand, the majority of reads that mapped to *A. brasilense* genome corresponded to unique mapped reads.

To identify genes differentially expressed in all libraries, statistical analyses using the R – DESeq2 package (Love *et al.* 2014) were performed. This package performs the multiple hypothesis testing to control the false discovery rate (FDR), and recently, it incorporates the

independent hypothesis weighting (IHW) into its analysis. The IHW is performed before the p-adjust be calculated, allowing an increase in the detection of differentially expressed genes (DEGs; Ignatiadis *et al.* 2016). To determine this increase in our data, we evaluated the amount of DEGs with and without applying the IHW approach. Comparing the number of DEGs detected when using just FDR with IHW+FDR methodology, we observed an increase of at least double in the detected DEGs (Table 3). Due to the remarkable increase in the detection power of the DEGs, we applied the IHW into our analysis.

Concerning RNA data for maize, the comparison between cDNA libraries from groups II, III, or IV with the cDNA library from the group I showed that several DEGs were down-regulated. For the other comparisons, the majority of the DEGS were up-regulated (Table S1). For *A. brasilense* RNA-seq analyses identified 14 DEGs, almost all of them up-regulated (Table S2).

Some of the DEGs detected for maize in all comparisons (Table S1) deserved to be highlighted (Table 4). Among genes involved in the response or synthesis of phytohormones, we identified DEGs that participate in response to ABA. Concerning these genes, three of them were differentially expressed: *PYL4* [Pyrabactin resistance 1 (*PYR1*)/*PYR1*-like (*PYL*)/regulatory components of ABA receptors (*RCAR*)] gene (Zm00001d028793), that codes for a cytoplasmatic ABA receptor, was up-regulated in two situations (group IV vs group I, and group IV vs group III); *RAE1* (for RNA export factor 1, Zm00001d025111), was strongly down-regulated in three situations (group II vs group I, group III vs group I, and group IV vs group I); and *ROP7* (Rho-related protein from plants 7, Zm00001d010732), which was also strongly down-regulated in two situations (group III vs I, and group IV vs I) and codes for the plant subfamily of the Rho small GTP-binding proteins. We also looked for the expression of the *YUCCA2* gene, since it was supposed to be the target for Yucasin effect. Even though it did not show statistically valid p-adjust values, its expression pattern in the cDNA libraries showed interesting results. When comparing RNA data from group II with the group I, *YUCCA2* expression was down-regulated. On the other hand, when comparing RNA data from group IV with those from groups II and III, the expression pattern of this gene was up-regulated (Table 4).

Other interesting DEGs were also observed: *WIN1* (Zm00001d048947), which is involved into the plant response to biotic and abiotic stresses, was down-regulated in the comparison between cDNA libraries from groups II and III with the group I (log2 Fold-change < - 4); *TYFY3B* gene (also known as *ZmTYFY*12 or ZmJaz8) (Zm00001d002029) was down-regulated in the comparison between cDNA library from group III with the group I; *MYB29* gene (MYB domain

protein 29) was strongly up-regulated (log2 Fold-change > 14) when comparing cDNA libraries from groups II and IV with the group I. Both genes (*TYFY3B* and *MYB29*) code for transcription factor (TF) proteins. Another TF family that is involved in the plant pathogen response is the WRKY family. In our study, the *WRKY 71* gene (Zm00001d006001) was up-regulated when comparing cDNA library from group IV with those from groups I, II, and III. This result could indicate that the plant was responding to the stress caused by the Yucasin and the presence of the bacterium. The leucine-rich repeat receptor-like kinase (LRR – RLK) *GASSHO* 1 (*GSO1*) was another gene with the expression reduced in the comparison between cDNA libraries from groups III and IV with that from group I.

Fifty-five uncharacterized loci were observed to be remarkably regulated (Table S3), and most of them were down-regulated. Among those is the uncharacterized locus LOC100285166 (Zm00001d021835), that is down-regulated when comparing the cDNA library from group IV with group II. This locus was identified as a calcium-dependent protein kinase (CPDK) that acts upstream of the RboH in the hypersensitive response pathway (data not shown).

Discussion

Azospirillum is a genus that contains several species with abilities to promote plant growth. Among their plant-growth promotion abilities, nitrogen fixation, and phytohormone secretion are the most studied so far (Thomas *et al.* 2019; Steenhoudt and Vanderleyden, 2000). Species of this genus were isolated from cereals and grasses and along the years the relationship between them and the plants have been studied by several researches (Holguin *et al.* 1999; Thomas *et al.* 2019; Ribaudo *et al.* 2006; Cassán, 2001; Cassán *et al.* 2009; Molina *et al.* 2018; Rivera *et al.* 2018). In the present work, the interaction between *Azospirillum brasilense* FP2 and maize during the inhibition of the maize's indole-3-acetic acid (IAA) production (TAA/YUC pathway) by Yucasin was investigated.

Since Yucasin did not interfere with the bacterial growth and IAA synthesis (Figures 1 and S1), the experiment to analyze the interaction between *Azospirillum brasilense* FP2 and maize during the inhibition of the maize's TAA/YUC pathway by Yucasin was started. Samples were taken five hours after the application of the Yucasin to analyze gene expression patterns of both organisms, and after five days of daily supplementation of this compound to verify its physiological effects to the plantlets. Plantlets from group II, which received Yucasin, presented a lower number of lateral roots than those from the other three groups (Figures 3 and S2). On the other hand, plantlets inoculated with *A. brasilense* (groups III and IV) showed no difference in the

number of lateral roots when compared with those from the group I (Figure 3), even in the presence of Yucasin (group IV). According to Lubna *et al.* (2018), application of Yucasin on maize seedling affected seedling growth, chlorophyll, and phenol contents. As IAA produced in root tips is one of the hormones involved in the formation of lateral roots (Jing and Strader, 2019) and Yucasin can prevents the conversion of IPA in IAA by YUC (Nishimura *et al.* 2014), we concluded that in fact Yucasin inhibited the formation of lateral roots of plantlets from group II. Lubna *et al.* (2018) also demonstrated that the production of IAA and other phytohormones by the fungus *Aspergillus niger* strain CSR3 could alleviate the effects of Yucasin on maize seedling. In our study, the number of lateral roots of plants inoculated with *A. brasilense* FP2 and under Yucasin treatment (group IV) was equivalent to those that did not receive Yucasin (group I) or were inoculated with the bacterium (group III; Figures 3 and S2). These results suggested that the bacterial PGP effects were able to recover the phenotype of plantlet's roots caused by the Yucasin (compare groups II and IV in Figure 3). Also, through an electron microscopy analysis, we were able to visualize colonies of *A. brasilense* in the maize's roots in both treatments, with and without Yucasin (Figure S3).

Once observed that Yucasin effectively inhibit lateral roots formation, and *A. brasilense* could reverse this phenotype and even promote root growth in the presence of Yucasin, we performed a dual RNA-seq analysis. All cDNA libraries were trimmed and mapped against the respective reference genome (*Z. mays* B73 genome for libraries of groups I and II, and a combined reference for libraries of group III and IV since these libraries contained reads from *Z. mays* and *A. brasilense*). Concerning the reads that mapped to the maize genome, most of the mapped ones corresponded to multi-reads. Since these reads aligned to CDS, and not to rRNA or tRNA, we concluded that the presence of multi-reads must be because *Z. mays* is an allopolyploid plant (Messing, 2009), presenting several copies of many genes. This fact corroborates Mortazavi *et al.* (2008) assertion that multi-reads should not be discarded under the penalty of losing transcriptional information. On the other hand, the majority of reads that mapped to *A. brasilense* genome corresponded to unique mapped reads. This can be explained by the well-known fact that the prokaryote genomes are formed by single-copy genes, exception for rRNA and tRNA genes.

RNA-seq analysis identified maize's differentially expressed genes that code for proteins involved in response to biotic and abiotic stresses, and synthesis or response to hormones (Table S1). Between those involved in response or synthesis of phytohormones were genes that participate in response to ABA and the synthesis and response to gibberellins (GA). About the plant's

response to ABA, three genes were differentially expressed. One of them was *PYL4*, which codes for a type of pyrabactin resistance 1 (*PYR1*)/*PYR1*-like (*PYL*)/regulatory components of ABA receptors (RCAR), that constitute a 14-member family (Bueso *et al.* 2014), and was up-regulated in two situations (Tables 4 and S1). In the presence of ABA, PYR/PYL receptors inhibit a clade A PP2Cs. This inhibition allows the activation of downstream targets of the PP2Cs, such as the sucrose non fermenting 1-related subfamily 2 (SnRK2) protein kinases. On their turn, these kinases act in the regulation of ABA signaling, including the regulation of transcriptional response to ABA and stomatal aperture (Pizzio *et al.* 2013; Bueso *et al.* 2014). Another gene was *RAE1*, which was strongly down-regulated in three situations (Tables 4 and S1). Li *et al.* (2018) suggested that, in *Arabidopsis*, the CUL4–DDB1 E3 ligase might directly target RCAR1 for ubiquitylation and degradation using the AtRAE1 (RNA export factor 1 in *Arabidopsis*) as the substrate receptor. According to them, AtRAE1 acts as a negative regulator of ABA signaling. The third gene identified in our RNA-seq analysis was *ROP7*, that codes for a Rho-related protein from plants 7 (Ivanchenko *et al.* 2000), and is part of the plant subfamily of the Rho small GTP binding proteins. These proteins function as molecular switches due to changes in conformation upon GTP binding and hydrolysis (Feiguelman *et al.* 2018). Feiguelman *et al*. (2018) postulated that ROP interacts with ABA-insensitive 1 and 2 (ABI1/2) PP2C phosphatases and prevents their inactivation by PYR/PYL ABA receptors. In the presence of ABA, ROP GEFs (guanine nucleotide exchange factors) are marked to be degraded, leading to the inhibition of ROP and PP2C sequestration by PYR1/PYL ABA receptors (Feiguelman *et al.* 2018). Taking the above scenario in consideration, we postulated that the presence of *Azospirillum* and Yucasin, either alone or simultaneously, induced a response of genes involved in the ABA pathway (Figure 4). According to our results, *RAE1* was repressed in the presence of Yucasin or the bacterium, with no difference between the repression levels caused by them individually or combined. These results suggested that the repression of *RAE1* expression was part of the plant's stress response mediated by ABA. In the treatment with only the bacterium, *RAE1* and *ROP7* were repressed. Since *ROP7* was strongly repressed, and both products of these genes respond to ABA levels, we can suppose that the presence of the bacterium somehow enhanced ABA response in the maize roots. The *RAE1* repression prevents PYR1/PYL to be marked for degradation, and *ROP* repression prevents it from protecting ABI1 PPC2 phosphatases from being inhibited by PYR1/PYL receptors. In summary, the presence of both bacterium and Yucasin led to the repression *RAE1*, to the up-regulation of *PYL4*, and to the down-regulation of *ROP7*. Since Feiguelman *et al*. (2018) postulated that

somehow auxin could interfere in the *ROP* expression, we suggest that auxin, together with ABA, suppress *ROP7* expression. Besides, as *A brasilense* produce ABA (Cohen *et al.* 2008; Cohen *et al.* 2009; Cohen *et al.* 2015), and auxin (Spaepen *et al.* 2007; Spaepen and Vanderleyden, 2010) the presence of the bacterium probably increased the plant's abiotic stress response through ABA response genes. Figure 4 shows a tentative model to represent the interference of *A. brasilense* and Yucasin in the plant's ABA response based on data from the present work.

Concerning the plant response to biotic and abiotic stresses, gene *WIN1* was down-regulated when cDNA libraries from groups II and III were compared with that from the group I (log2 Fold-change < -4). This gene codes for a protein that is part of Class II Pathogen-Response 4 (Class II PR-4) protein family (Bravo *et al.* 2003; Agostini *et al.* 2018). In rice, the expression of this gene is triggered by both abiotic and biotic stresses and has four copies (Wang *et al.* 2011). These authors evaluated the expression of all four copies and noticed that IAA increased the relative expression of them from 0 to 6 hours after hormone application. On the other hand, when ABA was administrated, the relative expression decreased in the first three hours and increased after six hours of treatment. It is well-know that ABA is a "stress" hormone involved in the adaptive response to abiotic stress and as *A brasilense* can produce ABA (Cohen *et al.* 2008; Cohen *et al.* 2009) or stimulates its production by the plant during the interaction (Cohen *et al.* 2015), we can hypothesize that the observed repression of *WIN1* expression in groups II and III plantlets was probably caused by the ABA produced by the plant in response to Yucasin and the bacterium.

Genes coding for transcription factors (TFs) were also differentially expressed. In the presence of *A. brasilense*, *TYFY 3B* gene (also known as ZmTYFY 12 or ZmJaz8; Zhang *et al.* 2015) was down-regulated (group III vs. I; Table 4). This gene is part of a TF family formed by 30 genes in *Z. mays*, and is involved in the response of several plant tissues to biotic and abiotic stresses (Zhang *et al.* 2015). Since the unique treatment between groups I and III is the bacterium and associated with the fact that *WIN1* was also down-regulated in this comparison, we suggest that the presence of *Azospirillum* probably was the reason for the reduction of *WIN1* and *TYFY 3B* expressions, as both genes are related to the response to biotic stress (Wang *et al.* 2011; Zhang *et al.* 2015; Trevisan *et al.* 2017).

When comparing cDNA libraries from groups II (Yucasin treatment) and IV (Yucasin and bacterium treatment) with that from the group I, *MYB29* gene (MYB domain protein 29) was strongly up-regulated (log2 Fold-change > 14). MYB29 belongs to a TF family involved in controlling many cellular processes, like biotic and abiotic stress responses, development,

differentiation, metabolism, and defense (Ambawat *et al.* 2013). Recently, MYB29 was characterized as a regulator of Met-derived aliphatic glucosinolate biosynthesis in *Arabidopsis*, involved in the defense against biotic feeding pests and abiotic stress (Zhang *et al.* 2017). Our data suggested that the up-regulated pattern of *MYB29* is the plant response to Yucasin presence.

Another family of TF that is involved into the plant pathogen response is the WRKY family (Agostini *et al.* 2018; Eulgem and Somssich, 2007). Agostini *et al.* (2018) demonstrated that WRKY genes were involved in the induced systemic response (ISR) in maize during the infection by *Fusarium verticillioides*, being up-regulated during the infection. In our study, *WRKY 71* was up-regulated in the comparisons between cDNA library from group IV with all other groups, indicating that the plant was responding to the stress caused by the Yucasin and to the presence of the bacterium.

The leucine-rich repeat receptor-like kinase (LRR – RLK) *GASSHO* 1 (*GSO1*) gene was down-regulated when cDNA libraries from groups III and IV were compared with that from group I. RLKs are involved in the regulation of various signaling pathways in plants, such as plant growth and development, and plant defense response (Osakabe *et al.* 2013; Thomas *et al.* 2019). Some RLKs are even involved in the perception of microbial signaling molecules that are important to both disease resistance and symbiosis (Thomas *et al.* 2019). In *Arabidopsis*, *GSO1* works together with *GSO2*, and the absence of both causes several defects, including a high degree of permeability of root's epidermis (Tsuwamoto *et al.* 2008). Considering the fact that *Azospirillum* can actively produce phytohormones (Cohen *et al.* 2008; Cohen *et al.* 2009; Cohen *et al.* 2015; De-Bashan *et al.* 2008; Bashan and de-Bashan, 2010; Salazar-Cerezo *et al.* 2018) that can be used by the plant, the *GSO1* repression observed in this work could be motivated by the presence of *A. brasilense* in the roots. This repression probably helped the exchange of substances among the bacterium and the plant.

Finally, 55 uncharacterized loci were remarkably regulated (Table S3), and most of these were down-regulated. Locus LOC100285166 (Gene ID: 100285166) was down-regulated in the comparison between cDNA library from group IV with that from group II. This locus mapped in the plant-pathogen interaction pathway and was identified as codifying a calcium-dependent protein kinase (CPDK) that acts upstream of the RboH in the hypersensitive response pathway (data not shown). Since in these libraries the only difference is the presence of the bacterium, as both groups received Yucasin, one can assume that this down-regulation was due to the presence of the bacterium.

Unlike other works that focus just in the transcriptome of one of the interacting organisms (Boscari *et al.* 2013; Hegedűs *et al.* 2009; Pankievicz *et al.* 2016; Verwaaijen *et al.* 2017), we used a combined methodology to assess both transcriptomes. In our study, 14 genes of *A. brasilense* were identified as being differentially expressed (Table S2). The identified genes are part of general metabolic pathways, indicating that the bacteria were metabolically active. Although none of them participated of pathways involved in the plant-microbe interactions, their detections were important due to the fact that RNA from the microorganism is in a lower amount than that of eukaryote, which difficult its analyses. Another fact that difficult the detection of reads from the microorganism's cDNA libraries is that rRNA accounts for more than 80% of the total RNA of the prokaryotic cellular content (Petrova *et al.* 2017). According to Petrova *et al.* (2017), the depletion of the prokaryote rRNA increases the detection of mRNA in single or dual-species cultures. This observation corroborates our findings (Table 2), in which the most abundant RNA present in the cDNA libraries from groups III and IV concerning *A. brasilense* was rRNA. The high amount of bacterial rRNA can be explained by the fact that during cDNA libraries construction preference was given for the plant's rRNA depletion. Since most of bacterial rRNA was still present in the libraries, it was sequenced, making the detection of bacterial mRNA sequences even more difficult. Due to this, we recommend for Dual RNA-seq studies that specific rRNA depletion methodologies must be used to increase the detection of mRNAs from both interacting organisms.

Materials and Methods

Bacterial strain and growth curve conditions

Azospirillum brasilense strain FP2 is a natural mutant originated from strain Sp7 (ATCC29145) that presents resistance to nalidixic acid and streptomycin antibiotics (Pedrosa and Yates, 1984). *A. brasilense* FP2 growth curves were obtained by inoculation of 2 mL of 24 h bacterial pre-culture in 250 mL Erlenmeyer's flask containing 100 mL of King B medium (Glickmann and Dessaux, 1995) supplemented or not with 50 µM of Yucasin [5-(4-chlorophenyl)-4H-1,2,4-triazole-3-thiol; register number CAS: 26028-65-9] to reach an initial OD_{600nm} of 0.02. Cultures were incubated in an orbital shaker at 30^{o}C and 120 rpm. Samplings were taken after 2, 4, 6, 26, 30, and 32 h of bacterial growth. To measure the bacterial IAA production, 2 mL of bacterial culture from each sampling time were centrifuged for 5 min at 14 Krpm. The supernatant was collected and mixed with the Salkowski reagent at the proportion of 1:2, respectively

(Glickmann and Dessaux, 1995). Bacterial IAA production was measured as described by Ambrosini *et al.* (2012).

Seed inoculation, experimental conditions, and physiological experiment

The bacterial suspension was prepared by growing *A. brasilense* FP2 in 30 mL of NFbHPN medium supplemented with 5 mg L^{-1} of malic acid (Pedrosa and Yates, 1984) in an orbital shaker (30ºC, 120 rpm) until an OD_{600} of 0.8 (Faleiro *et al.* 2013). Aliquots of 3 mL of the culture were centrifuged, and the pellets were suspended in NFb medium without nitrogen.

Zea mays (var. Santa Helena SHS4080) seeds were surface-sterilized by washing them three times with autoclaved ultrapure water, followed by submersion in 70% ethanol for 3 min and in a solution of 2% sodium hypochlorite and 2.5% Tween 20 for 30 min. Seeds were then washed three times with sterile distilled water by gentle shaking (Faleiro *et al.* 2013).

The experiment was divided into four groups: groups I and II were formed by plantlets which seeds were not inoculated with *A. brasilense* FP2; groups III and IV were formed by plantlets which seeds were inoculated with *A. brasilense* FP2. To inoculation, seeds were mixed with a bacterial suspension of 0.05 mL $seed^{-1}$(Hungria *et al.* 2010) and incubated for 5 min in an orbital shaker at 100 rpm (Faleiro *et al.* 2013). Seeds inoculated or not were placed in a sterilized water-saturated paper and maintained for three days in a 25ºC growth chamber in the dark for germination. Maize seedlings were then transferred to pots containing sterilized sand wet with plant medium solution without nitrogen (Egener, 1999; Espindula *et al.* 2017). The plantlets were kept in a growing chamber for 10 days (25ºC, 16 h light/8 h dark, with active photosynthetic radiation of 150 µmol m^2s^{-1}) (Espindula *et al.* 2017; Faleiro *et al.* 2013). After this period, plants from groups II and IV received 50 µM of Yucasin diluted in sterilized water. Plants from groups I and III received only sterilized water. The experiment was carried out with three biological replicates. Each replicate consisted of 20 *Z. mays* plantlets. After five hours under Yucasin treatment roots were washed twice with sterilized water, separated from the aerial part and stored at -80ºC until RNA extraction and electron microscopy analysis.

To access the physiological effects of the Yucasin into the root's development, another four groups of maize plantlets were prepared as previously described. Ten days plantlets from groups II and IV were then daily supplemented with a 50 µM of Yucasin for additional five days, while plantlets from groups I and III received only sterilized water (Nishimura *et al.* 2014). After this period, the lengths of the aerial parts and roots and the number of lateral roots of each plantlet were evaluated.

Scanning Electron Microscopy (SEM)

To visualize *A. brasilense* cells attached to maize roots, two *Zea mays* roots from groups III and IV treatments were fixed with Karnovsky's fixative (Karnovsky, 1964) and washed in an alcoholic series (20, 40, 60,70, 80, 90, 96, and 100%) during 30 min at each concentration. After complete dehydration, the samples were dried at the critical point of CO_2 at the equipment Leica EM CDP 300 (Haddad *et al.* 2007)**.** The dried samples were attached on aluminum supports with the aid of double-sided carbon tape and were carbon-coated at Baltec's Sputter Coater, model CED 005, for ultrastructural study. The analysis was conducted in a Scanning Electron Microscope FEI, model Inspect F50, at the Central Laboratory for Microscopy and Microanalysis, Pontifical Catholic University of Rio Grande do Sul (LabCEMM-PUC/RS).

RNA isolation, mRNA enrichment, cDNA synthesis, and sequencing

Total RNA was isolated from 0.1 g of plant tissue from a pool of 20 maize roots for each biological replicate from each treatment. There were three RNA extractions per experimental group, one per biological replicate, totalizing 12 RNA samples. Total RNA was isolated by RNeasy Plant Mini Kit® (Qiagen, CA, USA). The RNA concentration and purity were determined by spectrophotometry at 260 nm and 280 nm (Jahn *et al.* 2008) measured in a Nanodrop LITE spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Then, samples were first treated with DNaseI (Invitrogen), and the rRNA pool was depleted using the RiboMinus™ Plant Kit for RNA-Seq (Invitrogen). The cDNA libraries were constructed using the Ion total RNA-Seq kit v2 for Whole Transcriptome Library. All RNA quantification and quality evaluation were performed at the Bioanalyzer™ - Agilent 2100 instrument. Each cDNA library obtained was sequenced using the Ion PI Template OT2 200 Kit v3 and the Ion PI Sequencing 200 Kit v3 at the IonTorrent® platform (Thermo FisherScientific, Wilmington, DE, USA). All kits and reagents were used according to manufacturer's instructions. The 12 cDNA libraries obtained in this work were deposited in GenBank under numbers SAMN12391479 to SAMN12391490.

Data Analysis and Differential Gene Expression

The reference genomes and the respective annotations were downloaded from the National Center for Biotechnology Information (NCBI) site. All the cDNA libraries obtained and the reference genomes with annotations were uploaded into the CLC Genomics Workbench (v. 8.0). Reads smaller than 20 nucleotides and with low quality were removed from libraries using the standards setup of CLC Genomics. Groups I and II cDNA libraries were mapped against the *Z. mays* cv. B73 (GCF_000005005.2) genome and the groups III and IV cDNA libraries were mapped

against a Combined reference file formed by the merging of *Z. mays* cv. B73 and *A. brasilense* Sp7 (GCA_001315015.1) genomes. The mapping parameter used was 80% alignment to the reference sequence and 80% identity required for inclusion as a mapped read. The mapped reads were extracted and counted using the respective annotated genome. The counting parameters used were: 80% alignment to the reference sequence and 80% identity required for inclusion as a mapped read, number of hits equal to 10, with the exclusion of the reads that mapped to intergenic regions (Mortazavi *et al.* 2008, with modofications).

Count files were analyzed with the DESeq2 v3.8 (Love *et al.* 2014) and IHW v3.8 (Ignatiadis *et al.* 2016) packages of R software v3.5.2 (R Development Core Team). For maize, differentially expressed genes (DEGs) were considered those that presented *p*-adjust< 0.1 and Log2 fold-change ≥ 2 (Schurch *et al.* 2016; Camilios-Neto *et al.* 2014). For *A. brasilense*, DEGs were considered those that presented *p*-value < 0.1 and Log2 fold-change ≥ 2.

Acknowledgments

This work was supported by the Brazilian funding agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), and by Newton Fund (Brazil-UK collaboration).

A special thanks to the professor PhD Sidia Maria Callegari Jacques for all the priceless advices and help during the statistical analysis of our data.

References

Agostini, R. B., Postigo, A., Rius, S. P., Rech, G. E., Campos-Bermudez, V. A. and Vargas, W. A. (2018) Long-lasting primed state in maize plants: salicylic acid and steroid signaling pathways as key players in the early activation of immune responses in silks, Mol. Plant Microbe In., 32(1), pp. 95-106.

Ambawat, S., Sharma, P., Yadav, N. R. and Yadav, R. C. (2013) MYB transcription factor genes as regulators for plant responses: an overview, Physiol Mol Biol Plants, 19(3), pp. 307-321.

Ambrosini, A., Beneduzi, A., Stefanski, T., Pinheiro, F. G., Vargas, L. K. and Passaglia, L. M. P. (2012) Screening of plant growth promoting rhizobacteria isolated from sunflower (*Helianthus annuus* L.), Plant Soil, 356(1), pp. 245-264.

Babalola, O. O. (2010) Beneficial bacteria of agricultural importance, Biotechnol Lett, 32(11), pp. 1559-70.

Bashan, Y. and de-Bashan, L. E. (2010) How the plant growth-promoting bacterium *Azospirillum* promotes plant growth—A Critical Assessment, Adv. Agron., pp. 77-136.

Baudoin, E. L., A., Mirza, M. S., El Zemrany, H.;Prigent-Combaret, C., Jurkevich, E., Spaepen, S., Vanderleyden, J., Nazaret, S., Okon, Y., Moenne-Loccoz, Y. (2010) Effects of *Azospirillum brasilense* with genetically modified auxin biosynthesis gene *ipdC* upon the diversity of the indigenous microbiota of the wheat rhizosphere, Res. Microbiol., 161(3), pp. 219-26.

Bishopp, A., Mahonen, A. P. and Helariutta, Y. (2006) Signs of change: hormone receptors that regulate plant development, Development, 133(10), pp. 1857-69.

Boscari, A., Del Giudice, J., Ferrarini, A., Venturini, L., Zaffini, A. L., Delledonne, M. and Puppo, A. (2013) Expression dynamics of the *Medicago truncatula* transcriptome during the symbiotic interaction with *Sinorhizobium meliloti:* which role for nitric oxide?, Plant Physiol, 161(1), pp. 425-39.

Bravo, J. M., Campo, S., Murillo, I., Coca, M. and San Segundo, B. (2003) Fungus- and wound-induced accumulation of mRNA containing a class II chitinase of the pathogenesis-related protein 4 (PR-4) family of maize, Plant Mol Biol, 52(4), pp. 745-759.

Bueso, E., Rodriguez, L., Lorenzo-Orts, L., Gonzalez-Guzman, M., Sayas, E., Muñoz-Bertomeu, J., Ibañez, C., Serrano, R. and Rodriguez, P. L. (2014) The single-subunit RING-type E3 ubiquitin ligase RSL1 targets PYL4 and PYR1 ABA receptors in plasma membrane to modulate abscisic acid signaling, Plant J, 80(6), pp. 1057-1071.

Camilios-Neto, D., Bonato, P., Wassem, R., Tadra-Sfeir, M. Z., Brusamarello-Santos, L. C. C., Valdameri, G., Donatti, L., Faoro, H., Weiss, V. A., Chubatsu, L. S., Pedrosa, F. O. and Souza, E. M. (2014) Dual RNA-seq transcriptional analysis of wheat roots colonized by *Azospirillum brasilense* reveals up-regulation of nutrient acquisition and cell cycle genes, BMC Genomics, 15.

Cassán, F., Perrig, D., Sgroy, V., Masciarelli, O., Penna, C. and Luna, V. (2009) *Azospirillum brasilense* Az39 and *Bradyrhizobium japonicum* E109, inoculated singly or in combination, promote seed germination and early seedling growth in corn (*Zea mays* L.) and soybean (*Glycine max* L.), Eur J Soil Biol, 45(1), pp. 28-35.

Cassán, F. B., R.;Schneider, G.;Piccoli, P. (2001) *Azospirillum brasilense* and *Azospirillum lipoferum* hydrolyze conjugates of GA20 and metabolize the resultant aglycones to GA1 in seedlings of rice dwarf mutants1, Plant Physiol, 125, pp. 2053–2058.

Cohen, A. C., Bottini, R. and Piccoli, P. N. (2008) *Azospirillum brasilense* Sp 245 produces ABA in chemically-defined culture medium and increases ABA content in *Arabidopsis* plants, Plant Growth Regul, 54(2), pp. 97-103.

Cohen, A. C., Bottini, R., Pontin, M., Berli, F. J., Moreno, D., Boccanlandro, H., Travaglia, C. N. and Piccoli, P. N. (2015) *Azospirillum brasilense* ameliorates the response of *Arabidopsis thaliana* to drought mainly via enhancement of ABA levels, Physiol Plantarum, 153(1), pp. 79-90.

Cohen, A. C., Travaglia, C. N., Bottini, R. and Piccoli, P. N. (2009) Participation of abscisic acid and gibberellins produced by endophytic *Azospirillum* in the alleviation of drought effects in maize, Botany, 87(5), pp. 455-462.

De-Bashan, L. E., Antoun, H. and Bashan, Y. (2008) Involvement of indole-3-acetic acid produced by the growth-promoting bacterium *Azospirillum* spp. in promoting growth of *Chlorella vulgaris* 1, J Phycol, 44(4), pp. 938-947.

Duca, D., Lorv, J., Patten, C. L., Rose, D. and Glick, B. R. (2014) Indole-3-acetic acid in plant–microbe interactions, Antonie van Leeuwenhoek, 106, pp. 85-125.

Egener, T. H., T.;Reinhold-Hurek,B. (1999) Endophytic expression of *nif* genes of *Azoarcus* sp Strain BH72 in rice roots, Mol Plant Microbe In, 12(9), pp. 813-819.

Espindula, E., Faleiro, A. C., Pereira, T. P., Amaral, F. P. and Arisi, A. C. M. (2017) *Azospirillum brasilense* FP2 modulates respiratory burst oxidase gene expression in maize seedlings, Indian J Plant Physi, 22(3), pp. 316–323.

Eulgem, T. and Somssich, I. E. (2007) Networks of WRKY transcription factors in defense signaling, Curr Opin Plant Biol, 10(4), pp. 366-371.

Faleiro, A. L., Pereira, T. P., Espindula, E., Brod, F. C. A. and Arisi, A. C. M. 2013. Real time PCR detection targeting *nifA* gene of plant growth promoting bacteria *Azospirillum brasilense* strain FP2 in maize roots. Symbiosis, 61(3), pp. 125-133.

Feiguelman, G., Fu, Y. and Yalovsky, S. (2018) ROP GTPases structure-function and signaling pathways, Plant Physiol, 176(1), pp. 57-79.

Fornasieri Filho, D. (1992) *A Cultura do Milho.*Jaboticabal: FUNEP.

Glickmann, E. and Dessaux, Y. (1995) A critical examination of the specificity of the Salkowski reagent for indolic compounds produced by phytopathogenic bacteria, Appl Environ Microb, 61(2), pp. 793-796.

Haddad, A., Campos, A.P.C., Sesso and A. (2007) *Técnicas de Microscopia Eletrônica Aplicadas às Ciências Biológicas.* 3 edn. Rio de Janeiro, Sociedade Brasileira de Microscopia.

Hegedűs, Z., Zakrzewska, A., Ágoston, V. C., Ordas, A., Rácz, P., Mink, M., Spaink, H. P. and Meijer, A. H. (2009) Deep sequencing of the zebrafish transcriptome in response to mycobacterium infection, Mol Immunol, 46(15), pp. 2918-2930.

Holguin, G., Patten, C. L. and Glick, B. R. (1999) Genetics and Molecular Biology of *Azospirillum*, Biol Fert Soils, 29, pp. 10-23.

Hungria, M., Campo, R. J., Souza, E. M. and Pedrosa, F. O. (2010) Inoculation with selected strains of *Azospirillum brasilense* and *A. lipoferum* improves yields of maize and wheat in Brazil,Plant Soil, 331(1-2), pp. 413-425.

Ignatiadis, N., Klaus, B., Zaugg, J. B. and Huber, W. (2016) Data-driven hypothesis weighting increases detection power in genome-scale multiple testing, Nat Methods, 13, pp. 577.

Ivanchenko, M., Vejlupkova, Z., Quatrano, R. S. and Fowler, J. E. (2000) Maize ROP7 GTPase contains a unique CaaX box-independent plasma membrane targeting signal, Plant J, 24(1), pp. 79-90.

Jahn, C. E., Charkowski, A. O. and Willis, D. K. (2008) Evaluation of isolation methods and RNA integrity for bacterial RNA quantitation,J Microbiol Meth, 75(2), pp. 318-324.

Jing, H. and Strader, C. L. (2019) Interplay of Auxin and Cytokinin in Lateral Root Development, Int J Mol Sci, 20(3), pp. 486-498.

Karnovsky, M. (1964) A Formaldehyde-Glutaraldehyde Fixative of High Osmolality for Use in Electron Microscopy, J Cell Biol, 27(2), pp.1A-149A.

Li, D., Zhang, L., Li, X., Kong, X., Wang, X., Li, Y., Liu, Z., Wang, J., Li, X. and Yang, Y. (2018) AtRAE1 is involved in degradation of ABA receptor RCAR1 and negatively regulates ABA signalling in *Arabidopsis*, Plant Cell Environ, 41(1), pp. 231-244.

Li, Y., Wang, W., Feng, Y., Tu, M., Wittich, P. E., Bate, N. J. and Messing, J. (2019) Transcriptome and metabolome reveal distinct carbon allocation patterns during internode sugar accumulation in different sorghum genotypes, Plant Biotechnol J, 17(2), pp. 472-487.

Love, M. I., Huber, W. and Anders, S. (2014) Moderated estimation of fold-change and dispersion for RNA-seq data with DESeq2, Genome Biol, 15(12), pp. 550.

Lubna, Asaf, S., Hamayun, M., Gul, H., Lee, I.-J. and Hussain, A. (2018) *Aspergillusniger* CSR3 regulates plant endogenous hormones and secondary metabolites by producing gibberellins and indoleacetic acid, J Plant Interact, 13(1), pp. 100-111.

Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., Natsume, M., Hanada, A., Yaeno, T., Shirasu, K., Yao, H., McSteen, P., Zhao, Y., Hayashi, K.-i., Kamiya, Y. and Kasahara, H. (2011) The main auxin biosynthesis pathway in *Arabidopsis*, P Natl Acad Sci USA, 108(45), pp. 18512-18517.

Messing, J. (2009) The polyploid origin of maize, in Bennetzen, J.L. and Hake, S. (eds.) *Handbook of Maize: Genetics and Genomics*. New York, NY: Springer New York, pp. 221-238.

Molina, R., Rivera, D., Mora, V., López, G., Rosas, S., Spaepen, S., Vanderleyden, J. and Cassán, F. (2018) Regulation of IAA biosynthesis in *Azospirillumbrasilense* under environmental stress conditions, Curr Microbiol, 75(10), pp. 1408-1418.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq, Nat Methods, 5, pp. 621.

Nishimura, T., Hayashi, K., Suzuki, H., Gyohda, A., Takaoka, C., Sakaguchi1, Y., Matsumoto, S., Kasahara, H., Sakai, T., Kato, J., Kamiya, Y. and Koshiba1, T. (2014) Yucasin is a potent inhibitor of YUCCA, a key enzyme in auxin biosynthesis, Plant J, 77, pp. 352–366.

Osakabe, Y., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.-S. P. (2013) Sensing the environment: key roles of membrane-localized kinases in plant perception and response to abiotic stress, J Exp Bot, 64(2), pp. 445-458.

Pankievicz, V. C. S., Camilios-Neto, D., Bonato, P., Balsanelli, E., Tadra-Sfeir, M. Z., Faoro, H., Chubatsu, L. S., Donatti, L., Wajnberg, G., Passetti, F., Monteiro, R. A., Pedrosa, F. O. and Souza, E. M. (2016) RNA-seq transcriptional profiling of *Herbaspirillum seropedicae* colonizing wheat (*Triticum aestivum*) roots, Plant Mol Biol, 90(6), pp. 589-603.

Pedrosa, F. O. and Yates, M. G. (1984) Regulation of nitrogen fixation (*nif*) genes of *Azospirillum brasilense* by *nifA* and *ntr* (*gln*) type gene products, FEMS Microbiol Lett, 23(1), pp. 95-101.

Petrova, O. E., Garcia-Alcalde, F., Zampaloni, C. and Sauer, K. (2017) Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes, Sci Rep - UK, 7, pp. 41114.

Perrig, D., Boiero, M. L., Masciarelli, O. A., Penna, C., Ruiz, O. A., Cassan, F. D. and Luna, M. V. (2007) Plant-growth-promoting compounds produced by two agronomically important strains of *Azospirillum brasilense*, and implications for inoculant formulation, Appl Microbiol Biot, 75(5), pp. 1143-50.

Pizzio, G. A., Rodriguez, L., Antoni, R., Gonzalez-Guzman, M., Yunta, C., Merilo, E., Kollist, H., Albert, A. and Rodriguez, P. L. (2013) The PYL4 A194T Mutant Uncovers a Key Role of

PYR1-Like 4/Protein phosphatase 2CA interaction for abscisic acid signaling and plant drought resistance, Plant Physiol, 163(1), pp. 441-455.

Ribaudo, C. M., Krumpholz, E. M., Cassán, F. D., Bottini, R., Cantore, M. L. and Curá, J. A. (2006) *Azospirillum* sp. promotes root hair development in tomato plants through a mechanism that involves ethylene, J Plant Growth Regul, 25(2), pp. 175-185.

Rivera, D., Mora, V., Lopez, G., Rosas, S., Spaepen, S., Vanderleyden, J. and Cassan, F. (2018) New insights into indole-3-acetic acid metabolism in *Azospirillum brasilense*, J Appl Microbiol, DOI: 10.1111/jam. 14080.

Salazar-Cerezo, S., Martínez-Montiel, N., García-Sánchez, J., Pérez-y-Terrón, R. and Martínez-Contreras, R. D. (2018) Gibberellin biosynthesis and metabolism: A convergent route for plants, fungi and bacteria, Microbiol Res, 208, pp. 85-98.

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., Blaxter, M. and Barton, G. J. (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?, RNA, 22(6), pp. 839-51.

Spaepen, S. and Vanderleyden, J. (2010) Auxin and plant-microbe interactions, Cold SpringHarbour Perspective Biology, e3.a001438.

Spaepen, S., Vanderleyden, J. and Remans, R. (2007) Indole-3-acetic acid in microbial and microorganism-plant signaling, FEMS Microbiol Rev, 31(4), pp. 425-448.

Steenhoudt, O. and Vanderleyden, J. (2000) *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: Genetic, biochemical and ecological aspects, FEMS Microbiol Rev, 24, pp. 487-506.

Thomas, J., Kim, H. R., Rahmatallah, Y., Wiggins, G., Yang, Q., Singh, R., Glazko, G. and Mukherjee, A. (2019) RNA-seq reveals differentially expressed genes in rice (*Oryza sativa*) roots during interactions with plant-growth promoting bacteria, *Azospirillum brasilense*, Plos One, 14(5), pp. e0217309.

Trevisan, S., Manoli, A., Ravazzolo, L., Franceschi, C. and Quaggiotti, S. (2017) mRNA-Sequencing analysis reveals transcriptional changes in root of maize seedlings treated with two increasing concentrations of a new biostimulant, J Agr Food Chem, 65(46), pp. 9956-9969.

Tsuwamoto, R., Fukuoka, H. and Takahata, Y. (2008) *GASSHO1* and *GASSHO*2 encoding a putative leucine-rich repeat transmembrane-type receptor kinase are essential for the normal development of the epidermal surface in *Arabidopsis* embryos, Plant J, 54(1), pp. 30-42.

Verna, J.P., Yadav, J., Tiwari, K. N., Singh, L. and Singh, V. (2010) Impact of plant growth promoting rhizobacteria on crop production, Int J Agric Res, 5(11), pp. 954-983.

Verwaaijen, B., Wibberg, D., Kröber, M., Winkler, A., Zrenner, R., Bednarz, H., Niehaus, K., Grosch, R., Pühler, A. and Schlüter, A. (2017) The *Rhizoctonia solani* AG1-IB (isolate 7/3/14) transcriptome during interaction with the host plant lettuce (*Lactuca sativa* L.), PLOS One, 12(5), e0177278.

Wang, N., Xiao, B. and Xiong, L. (2011) Identification of a cluster of PR4-like genes involved in stress responses in rice,J Plant Physiol, 168(18), pp. 2212-2224.

Won, C., Shen, X., Mashiguchi, K., Zheng, Z., Dai, X., Cheng, Y., Kasahara, H., Kamiya, Y., Chory, J. and Zhao, Y. (2011) Conversion of tryptophan to indole-3-acetic acid by Tryptophan Aminotransferases of *Arabidopsis* and YUCCAs in *Arabidopsis*, P Natl Acad Sci USA, 108(45), pp. 18518-18523.

Yilmaz, A., Nishiyama, M. Y., Fuentes, B. G., Souza, G. M., Janies, D., Gray, J. and Grotewold, E. (2009) GRASSIUS: A Platform for comparative regulatory genomics across the grasses, Plant Physiol, 149(1), pp. 171-180.

Yue, J., Hu, X. and Huang, J. (2014) Origin of plant auxin biosynthesis, Trends Plant Sci, 19(12), pp. 764-770.

Zhang, X., Ivanova, A., Vandepoele, K., Radomiljac, J., Van de Velde, J., Berkowitz, O., Willems, P., Xu, Y., Ng, S., Van Aken, O., Duncan, O., Zhang, B., Storme, V., Chan, K. X., Vaneechoutte, D., Pogson, B. J., Van Breusegem, F., Whelan, J. and De Clercq, I. (2017) The transcription factor MYB29 is a regulator of Alternative oxidase 1a, Plant Physiol, 173(3), pp. 1824-1843.

Zhang, Z., Li, X., Yu, R., Han, M. and Wu, Z. (2015) Isolation, structural analysis, and expression characteristics of the maize TIFY gene family, Mol Genet Genomics, 290(5), pp. 1849-1858.

Zhao, Y. (2010) Auxin biosynthesis and its role in plant development, Annu Rev Plant Biol, 61, pp. 49-64.

Zhao, Y. (2012) Auxin biosynthesis: A simple two-step pathway converts tryptophan to indole-3-acetic acid in plants, MolecularPlant, 5(2), pp. 334-338.

Zhao, Y. (2014) Auxin biosynthesis, The *Arabidopsis* book, pp. e0173.

Figures and Tables Legends

Figure 1. Amount of IAA produced over time by *Azospirillum brasilense* FP2 in King B medium supplemented with tryptophan with or without Yucasin.

Figure 2. Lengths of the roots (A) and the aerial parts (B) of maize plantlets according to each experimental group (see Material and Methods for group composition). Bars show twice standard error. DAI = days after inoculation.

Figure 3. Number of lateral roots of plantlets 15 days after inoculation (DAI, for groups III and IV), and after five days with Yucasin treatment (groups II and IV). Bars show twice standard error.

Figure 4. Model to represent the interference of *A. brasilense* and Yucasin in the plant's ABA response. *RAE1* was down-regulated when comparing gene expression of groups II, III, and IV with group I. *PYL4* was up-regulated when comparing gene expression of group IV with those of groups I and III. *ROP7* was down-regulated when comparing gene expression of groups III and IV with group I. Blank indicates genes not detected. Gray lines indicate reactions that probably are not occurring due to the expression of the detected genes. Probably the ABA produced by the bacterium or its increase in plant's production due to the presence of the bacterium and the Yucasin leads *RAE1* to be downregulated. Since PYL4 is not being marked for denaturation by RAE1, it is free to interact with ABA and inhibits PP2Cs activities. In consequence, SnRK2s are free to stimulate the expression of genes involved in the ABA response. On the other hand, ABA alongside with auxins produced by the bacterium, could be responsible for the inhibition observed for *ROP7*. ROP is known to interacts with PP2Cs, preventing them to be inhibited by PYR/PYL. All together we can suggest that the presence of the bacterium increased the plant abiotic stress response. SnRK2s = sucrose non fermenting 1-related subfamily 2 (SnRK2) protein kinases; PP2Cs = clade A protein phosphatases type-2C.

Table 1. Library features and number of total reads attributed to *Zea mays* or to the combined reference.

Table 2. Number of reads mapped to tRNA, rRNA, and coding sequences (CDS) in each experimental group.

Table 3. Amount of maize Differential Expressed Genes (DEGs) detected according the methodology used to calculate the p-adjusted. FDR: False Discovery Rate; IHW: Independent Hypothesis Weighting.

Table 4. Highlighted maize's DEGs in each experimental condition.

Supplementary Figures and Tables

Figure S1. Growth curves (OD_{600nm}) of *A. brasilense* FP2 in King B medium supplemented or not with Yucasin.

Figure S2. Plantlets from experimental groups I (A), II (B), III (C) and IV (D) at 15 DAI. Groups I and III were daily supplemented with autoclaved water and groups II and IV were daily supplemented with 50 µM of Yucasin for five days. Bars represent the scale, where each block has 1,0 cm.

Figure S3. Scanning electron microscopy of maize roots inoculated with *A. brasilense* strain FP2. **(A)** Roots from group III; **(B and C)** Roots from group IV. The pictures at **(C)** represent the same site with different magnifications. Arrows indicates *Azospirillum brasilense* FP2 colonies.

Table S1. Total maize's DEGs in each experimental condition.

Table S2. *Azospirillum brasilense* DEGs in each experimental condition.

Table S3. Maize's differential expressed uncharacterized genes in each experimental condition.

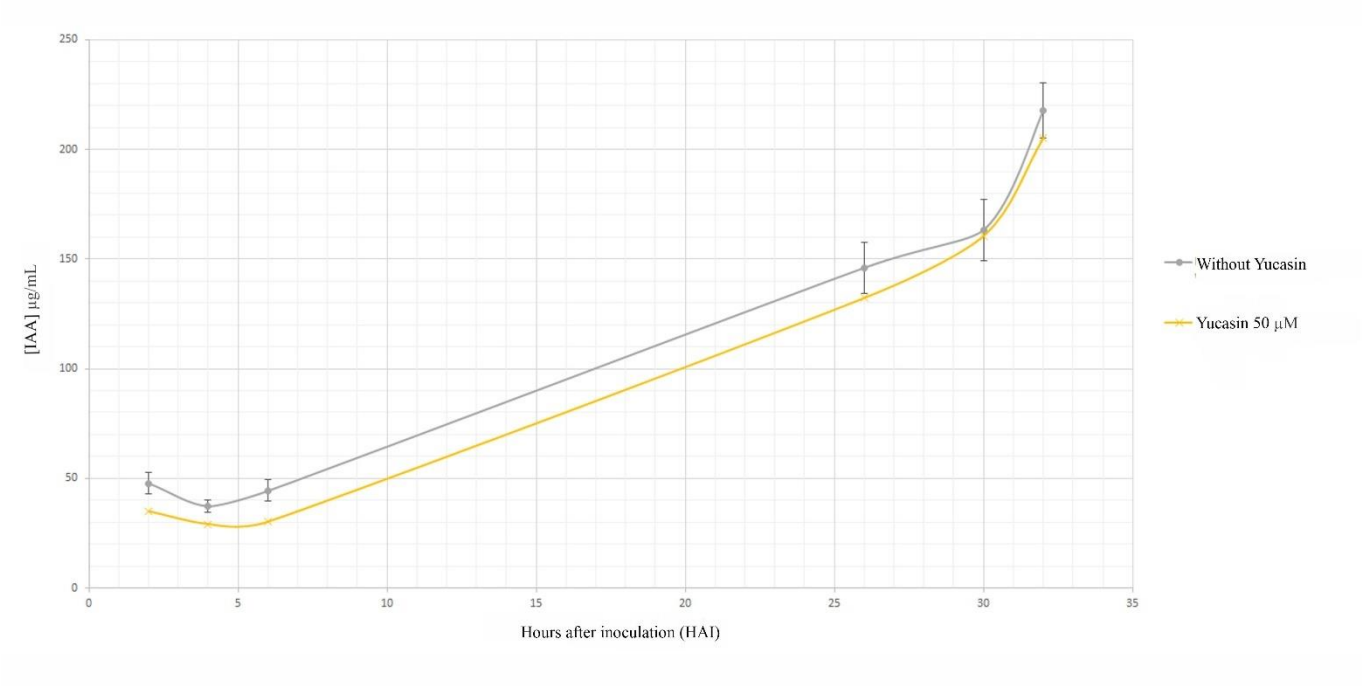


Figure 1. Amount of IAA produced over time by *Azospirillum brasilense* FP2 in King B medium supplemented with tryptophan with or without Yucasin.

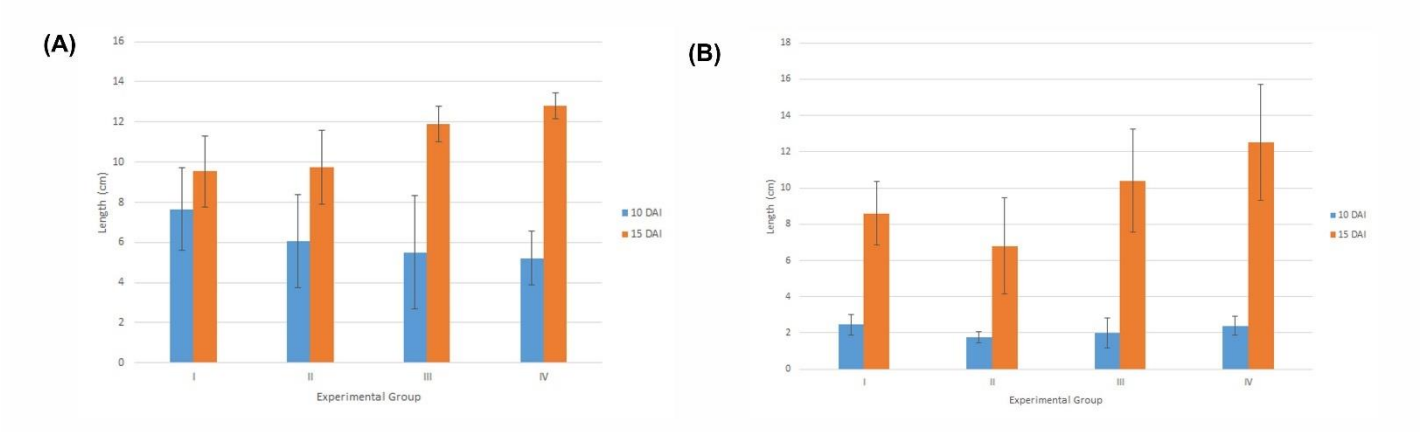


Figure 2. Lengths of the roots (A) and the aerial parts (B) of maize plantlets according to each experimental group (see Material and Methods for group composition). Bars show twice standard error. DAI = days after inoculation.

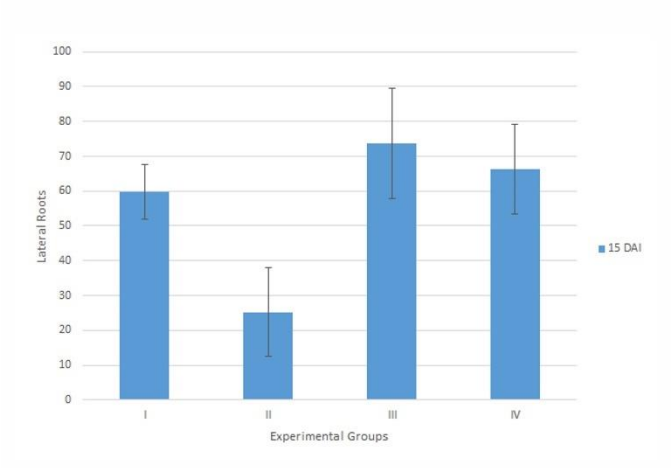


Figure 3. Number of lateral roots of plantlets 15 days after inoculation (DAI, for groups III and IV), and after five days with Yucasin treatment (groups II and IV). Bars show twice standard error.

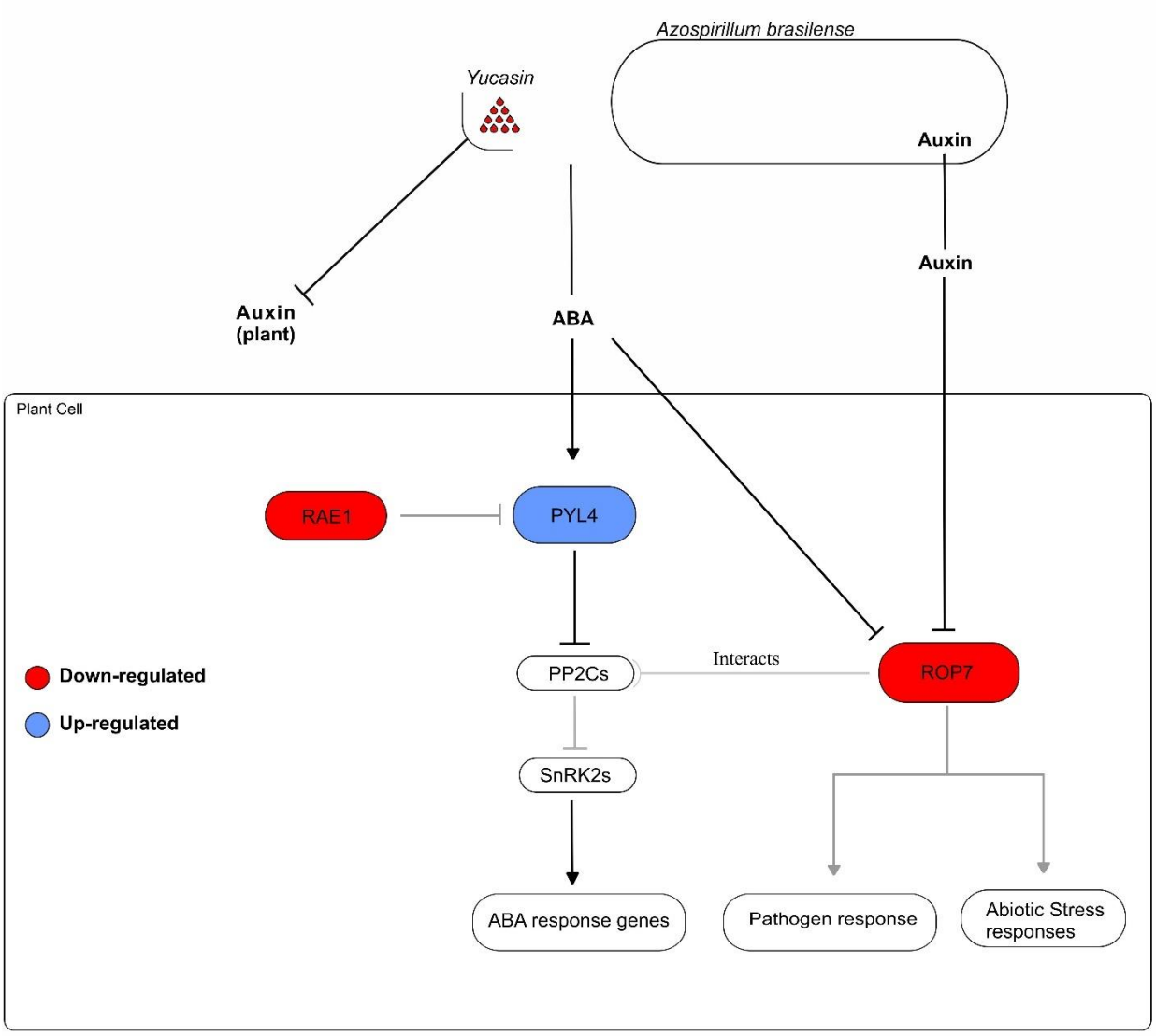


Figure 4. Model to represent the interference of *A. brasilense* and Yucasin in the plant's ABA response. *RAE1* was down-regulated when comparing gene expression of groups II, III, and IV with group I. *PYL4* was up-regulated when comparing gene expression of group IV with those of groups I and III. *ROP7* was down-regulated when comparing gene expression of groups III and IV with group I. Blank indicates genes not detected. Gray lines indicate reactions that probably are not occurring due to the expression of the detected genes. Probably the ABA produced by the bacterium or its increase in plant's production due to the presence of the bacterium and the Yucasin leads *RAE1* to be downregulated. Since PYL4 is not being marked for denaturation by RAE1, it is free to interact with ABA and inhibits PP2Cs activities. In consequence, SnRK2s are free to stimulate the expression of genes involved in the ABA response. On the other hand, ABA alongside with auxins produced by the bacterium, could be responsible for the inhibition observed for *ROP7*. ROP is known to interacts with PP2Cs, preventing them to be inhibited by PYR/PYL. All together we can suggest that the presence of the bacterium increased the plant abiotic stress response. SnRK2s = sucrose non fermenting 1-related subfamily 2 (SnRK2) protein kinases; PP2Cs = clade A protein phosphatases type-2C.

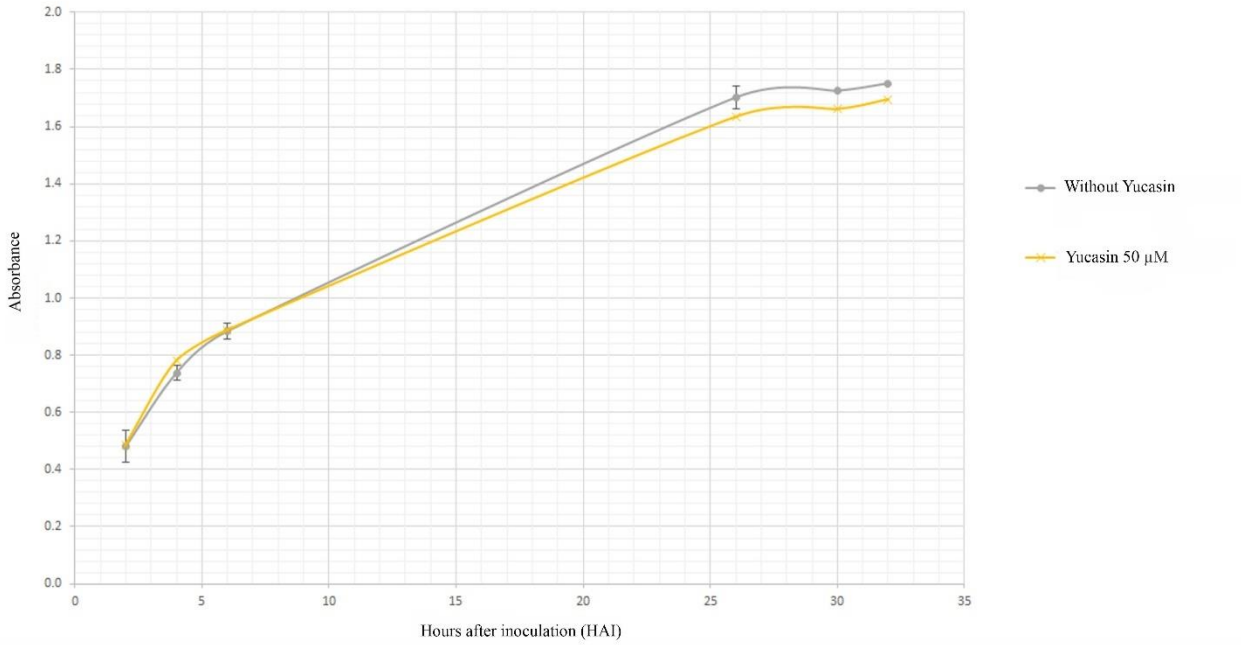


Figure S1. Growth curves (OD_{600nm}) of *A. brasilense* FP2 in King B medium supplemented or not with Yucasin.

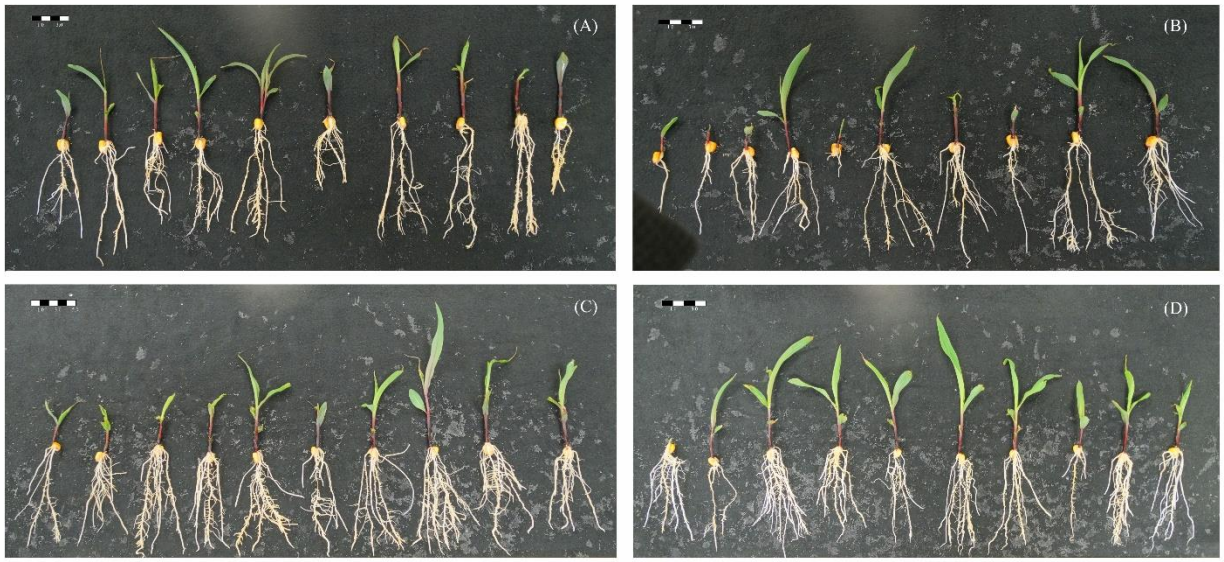


Figure S2. Plantlets from experimental groups I (A), II (B), III (C) and IV (D) at 15 DAI. Groups I and III were daily supplemented with autoclaved water and groups II and IV were daily supplemented with 50 µM of Yucasin for five days. Bars represent the scale, where each block has 1,0 cm.

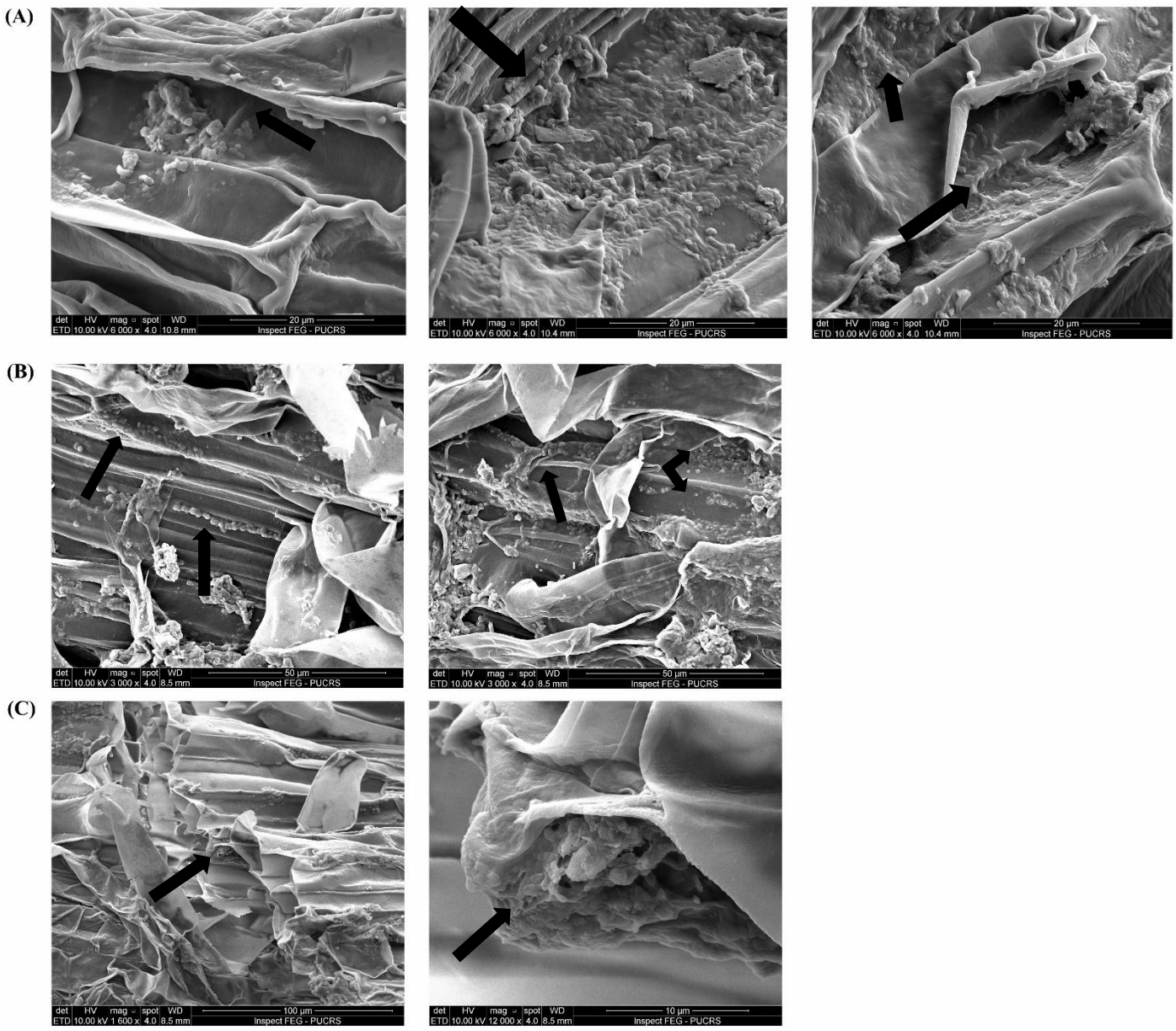


Figure S3. Scanning electron microscopy of maize roots inoculated with *A. brasilense* strain FP2. **(A)** Roots from group III; **(B and C)** Roots from group IV. The pictures at **(C)** represent the same site with different magnifications. Arrows indicates *Azospirillum brasilense* FP2 colonies.

Table 1. Library features and number of total reads attributed to *Zea mays* or to the combined reference.

Experimental Group	Total Reads per Experimental Group	Trimming (20 nt)	Reference used to Map the Reads	Reads Mapped to the reference	Percentage of Mapped Reads
I	51,749,063	49,793,940	*Zea mays*	45,111,393	91%
II	55,278,742	53,062,047	*Zea mays*	48,782,996	92%
III	57,905,272	55,537,430	Combined	50,232,328	91%
IV	50,309,689	46,047,663	Combined	40,288,886	88%

Table 2. Number of reads mapped to tRNA, rRNA, and coding sequences (CDS) in each experimental group.

Experimental Group	Library formed by the Reads that Mapped to	Total number of mapped reads	Number of Reads Mapped to				Unmapped Reads
			tRNA	rRNA	CDS sequences		
					Total Reads	Unique Reads	
I	*Zea mays*	45,111,393	11,826	1,949,984	24,786,788	810,125	18,362,795
II	*Zea mays*	48,782,996	7,736	2,022,472	29,210,896	1,043,121	17,541,892
III	*Azospirillum brasilense*	84,989	1,492	77,728	4,942	4,561	827
	Zea mays	50,147,339	10,483	2,538,964	25,016,612	1,267,707	22,581,280
IV	*Azospirillum brasilense*	106,413	1,753	94,841	8,420	7,762	1,399
	Zea mays	40,182,473	11,553	2,317,501	17,509,066	1,300,353	20,344,353

Table 3. Amount of maize Differential Expressed Genes (DEGs) detected according the methodology used to calculate the p-adjusted. FDR: False Discovery Rate; IHW: Independent Hypothesis Weighting.

Experimental Condition	Method for calculate the p-adjust	
	FDR	IHW + FDR
Group II vs Group I	7	24
Group III vs Group I	20	40
Group IV vs Group I	26	85
Group IV vs Group II	4	8
Group IV vs Group III	0	7

Table 4. Highlighted maize's DEGs in each experimental condition.

Experimental Condition	Locus tag	Gene ID	Description	log2 Fold-Change	p-value	p-adj	References
GII vs GI	Zm00001d025111	103641534	Protein RAE1 (RNA export factor 1)	-6.151233658	5.91E-07	0.005684345	(Li *et al.* 2018)
	Zm00001d048947	103652813	barwin-like (defense response to bacteria) (WIN1)	-4.321876905	1.52E-06	0.006851272	(Bravo *et al.* 2003)
	Zm00001d046517	103638687	transcription factor MYB29	14.86353597	3.87E-05	0.038420757	
	Zm00001d011764	100192827	/ Indole-3-pyruvate monooxygenase - YUCCA2 [*Zea mays*]	-1.24074068	0.361615987	1.00000000	
GIII vs GI	Zm00001d010732	542503	Rho-related protein from plants 7 (ROP7)	-9.305644274	2.05E-05	0.01105544	(Ivanchenko *et al.* 2000)
	Zm00001d025111	103641534	Protein RAE1 (RNA export factor 1)	-6.960927792	1.72E-08	5.43E-05	(Li *et al.* 2018)
	Zm00001d048947	103652813	barwin-like (defense response to bacteria) (WIN1)	-5.117430949	9.48E-08	0.000191566	(Bravo *et al.* 2003)
	Zm00001d037666	103630101	LRR receptor-like serine/threonine-protein kinase GSO1	-3.334088821	4.86E-05	0.014311343	(Tsuwamoto *et al.* 2008)
	Zm00001d002029	100384736	Transcription factor TIFY3B (aka: TIFY12)	-2.370662964	7.08E-06	0.005016479	(Zhang *et al.* 2015)
GIV vs GI	Zm00001d010732	542503	Rho-related protein from plants 7 (ROP7)	-7.266205411	0.000452026	0.057556	(Ivanchenko *et al.* 2000)
	Zm00001d025111	103641534	Protein RAE1 (RNA export factor 1)	-6.170564345	5.36E-07	0.0007	(Li *et al.* 2018)
	Zm00001d037666	103630101	LRR receptor-like serine/threonine-protein kinase GSO1	-2.972366806	0.000248077	0.03961	(Tsuwamoto *et al.* 2008)
	Zm00001d028793	103634514	Abscisic acid receptor PYL4 - like	2.386251179	0.001771926	0.098194	(Bueso *et al.* 2014)
	Zm00001d006001	100384128	WRKY transcription factor 71 - GRMZM2G052671 (Maize GDB)	2.483066281	0.000519928	0.049512	(Yilmaz *et al.* 2009)
	Zm00001d046517	103638687	transcription factor MYB29	20.98768324	2.75E-09	0.000007	

GIV vs GII	Zm00001d021835	100285166	uncharacterized LOC100285166	-3.204126327	8.77E-05	0.061676621	
	Zm00001d006001	100384128	WRKY transcription factor 71 - GRMZM2G052671 (Maize GDB)	2.957987635	3.79E-05	0.047599845	(Yilmaz *et al.* 2009)
	Zm00001d011764	100192827	Indole-3-pyruvate monooxygenase - YUCCA2 [*Zea mays*]	1.645618693	0.204677713	1.00000000	
GIV vs GIII	Zm00001d006001	100384128	WRKY transcription factor 71 - GRMZM2G052671 (Maize GDB)	3.319173225	4.66E-06	0.018467	(Yilmaz *et al.* 2009)
	Zm00001d028793	103634514	Abscisic acid receptor PYL4 - like	3.338990236	1.63E-05	0.028927	(Bueso *et al.* 2014)
	Zm00001d011764	100192827	Indole-3-pyruvate monooxygenase - YUCCA2 [*Zea mays*]	0.858712735	0.468954806	1.00000000	

Table S1. Total maize's DEGs in each experimental condition.

Experimental Condition	Gene ID	Description	log2 Fold-Change	p-value	p-adj
GII vs GI	103641534	Protein RAE1	-6.151233658	5.91E-07	0.005684345
	103646825	uncharacterized LOC103646825	-5.971302807	1.53E-05	0.038420757
	103630024	pentatricopeptide repeat-containing protein At5g48730, chloroplastic	-5.575428491	3.79E-05	0.032770133
	100280034	putative cytochrome P450 superfamily protein	-5.570676273	9.27E-06	0.02229954
	103644253	uncharacterized LOC103644253	-5.458778999	4.28E-05	0.038420757
	111590293	uncharacterized LOC111590293	-4.61713408	1.74E-05	0.038420757
	100283652	gibberellin receptor GID1L2	-4.530097368	0.000123944	0.05294829
	100283525	Protein MAK16 homolog (Uniprot - K7V9X2_MAIZE)	-4.482816438	8.68E-05	0.048842959
	100280286	uncharacterized LOC100280286	-4.448925606	9.46E-06	0.02229954
	103652813	barwin-like (defense response to bacteria) (win1)	-4.321876905	1.52E-06	0.006851272
	100217137	RNA pseudouridine synthase 7	-4.14650504	7.73E-05	0.048842959
	100194409	phosphatidate cytidylyl-transferase	-3.802533142	2.52E-05	0.025327815
	100272896	hypothetical protein	-3.730914165	0.00025149	0.053770186
	103636394	oxalate--CoA ligase	-3.692318776	9.86E-05	0.038420757
	103643524	ABC transporter G family member 25	-3.564983781	0.00018674	0.056384894
	100275449	uncharacterized LOC100275449	-3.541209705	4.33E-05	0.037875323
	100384583	Protein Chromatin remodeling 24	-3.137297272	0.000125836	0.077322515
	103655263	uncharacterized LOC103655263	-3.023127338	0.00015546	0.087565331
	103632291	phosphomannomutase/phosphoglucomutase	-2.840577576	2.31E-06	0.006851272
	100281356	ATP binding protein	-2.454107229	3.49E-05	0.037875323
	100381550	uncharacterized LOC100381550	-2.277649056	4.13E-05	0.032770133
	103644425	protein Shoot gravitropism 6	-2.035916334	6.29E-05	0.037875323
	100285433	AER (Transferase)	2.692665871	0.000256544	0.087565331
	103638687	transcription factor MYB29	14.86353597	3.87E-05	0.038420757

GIII vs GI	542503	Rho-related protein from plants 7	-9.305644274	2.05E-05	0.01105544
	103644253	uncharacterized LOC103644253	-8.771883724	1.05E-07	0.000340361
	100193510	U-box domain-containing protein 34	-7.80013243	7.67E-05	0.028537729
	103641534	Protein RAE1	-6.960927792	1.72E-08	5.43E-05
	103646825	uncharacterized LOC103646825	-6.381008475	4.59E-06	0.040656599
	103630024	pentatricopeptide repeat-containing protein At5g48730, chloroplastic	-6.340656703	3.59E-06	0.004627617
	100286099	pyrophosphate-energized vacuolar membrane proton pump	-5.949336031	0.000491555	0.06945181
	100279341	uncharacterized LOC100279341	-5.44058078	0.000221553	0.053829856
	109944009	uncharacterized LOC109944009	-5.347180688	2.05E-25	1.30E-21
	103652813	barwin-like (defense response to bacteria) (win1)	-5.117430949	9.48E-08	0.000191566
	100217137	RNA pseudouridine synthase 7	-4.306783298	3.86E-05	0.015245423
	100283525	Protein MAK16 homolog (Uniprot - K7V9X2_MAIZE)	-4.255190704	0.000176638	0.048547363
	100280286	uncharacterized LOC100280286	-4.119236592	3.01E-05	0.013078967
	100274041	uncharacterized LOC100274041	-4.008946848	0.000325405	0.077984566
	100277664	uncharacterized LOC100277664	-3.99506072	0.00013732	0.031031202
	100192499	uncharacterized LOC100192499	-3.982224712	0.000226684	0.067665291
	103654463	UDP-glycosyltransferase 89B2	-3.613142351	9.35E-06	0.005016479
	100194409	phosphatidate cytidylyl-transferase	-3.550993226	6.82E-05	0.029321177
	103630101	LRR receptor-like serine/threonine-protein kinase GSO1	-3.334088821	4.86E-05	0.014311343
	103636394	oxalate--CoA ligase	-3.270071316	0.000435414	0.090979988
	103643524	ABC transporter G family member 25	-3.262112871	0.000550949	0.086566236
	109940833	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic-like	-3.238237294	9.90E-08	0.000191566
	100384583	Protein Chromatin remodeling 24	-3.073422552	0.000139698	0.040656599
	100275449	uncharacterized LOC100275449	-2.966525509	0.000523379	0.08166411
	778429	low affinity nitrate transporter (ncbi)	-2.917068342	0.000292488	0.06945181
	542748	Thioesterase family protein, mRNA	-2.776119373	0.000174425	0.040656599

	103647807	formin-like protein 5	-2.77255045	0.00020085	0.062174438
	100284943	limonoid UDP-glucosyltransferase	-2.766552446	0.000503427	0.06945181
	103643201	phosphatidylinositol-3-phosphatase myotubularin-1	-2.658556975	0.000429795	0.06945181
	103632291	phosphomannomutase/phosphoglucomutase	-2.599083477	1.36E-05	0.009525088
	100381550	uncharacterized LOC100381550	-2.369286588	1.96E-05	0.011025741
	100384736	Transcription factor TIFY 3B (aka: TIFY12)	-2.370662964	7.08E-06	0.005016479
	100281939	60S ribosomal protein L29	-2.209648883	0.000373532	0.062174438
	103641772	ABC transporter B family member 15	-2.129246426	0.00090577	0.086086714
	100502224	uncharacterized LOC100502224	-2.131075988	1.36E-05	0.011332009
	100383448	putative RING zinc finger domain superfamily protein	2.174958628	0.00057321	0.076682608
	100272870	Acidic endochitinase	2.369573943	0.000658368	0.086566236
	606463	alkaline alpha galactosidase 1	2.42308656	2.36E-05	0.01105544
	100382848	uncharacterized LOC100382848	3.969145051	0.000180479	0.053829856
	100272285	CRS2-associated factor 1 mitochondrial	15.57942057	0.000280227	0.077489374
GIV vs GI	103644253	uncharacterized LOC103644253	-7.840435947	2.00E-06	0.001902
	542503	Rho-related protein from plants 7	-7.266205411	0.000452026	0.057556
	100193510	U-box domain-containing protein 34	-6.788669289	0.000353595	0.066717
	100384061	uncharacterized LOC100384061	-6.673609541	0.000854958	0.07479
	103628645	electron transfer flavoprotein subunit alpha, mitochondrial	-6.402765427	2.05E-05	0.018829
	103630024	pentatricopeptide repeat-containing protein At5g48730, chloroplastic	-6.402717599	2.87E-06	0.001599
	109944009	uncharacterized LOC109944009	-6.27415482	3.43E-32	4.52E-28
	103641534	Protein RAE1	-6.170564345	5.36E-07	0.0007
	100286099	pyrophosphate-energized vacuolar membrane proton pump	-6.015353811	0.000426779	0.059855
	103646825	uncharacterized LOC103646825	-5.592077697	7.33E-06	0.010616
	100274041	uncharacterized LOC100274041	-4.719772344	4.20E-05	0.010616
	100192499	uncharacterized LOC100192499	-4.451092579	5.05E-05	0.018829

111590293	uncharacterized LOC111590293	-4.3444928	2.07E-05	0.034623
100283525	Protein MAK16 homolog (Uniprot - K7V9X2_MAIZE)	-4.057986617	0.000332002	0.036186
103653702	uncharacterized LOC103653702	-3.979243115	0.000209593	0.060634
103636394	oxalate--CoA ligase	-3.703225136	8.15E-05	0.02303
100283652	gibberellin receptor GID1L2	-3.666890001	0.001186777	0.095106
100193544	uncharacterized LOC100193544	-3.523498625	0.000138305	0.028083
100272896	hypothetical protein	-3.49824169	0.000494378	0.059951
100280286	uncharacterized LOC100280286	-3.471764588	0.00035198	0.032714
103654463	UDP-glycosyltransferase 89B2	-3.46208719	1.49E-05	0.006743
542748	Thioesterase family protein, mRNA	-3.456871648	5.29E-06	0.005438
100381815	uncharacterized LOC100381815	-3.448432964	0.000279114	0.046466
778429	low affinity nitrate transporter (ncbi)	-3.436994045	2.78E-05	0.010616
100383495	uncharacterized LOC100383495	-3.374979564	0.000169831	0.035184
109940833	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic-like	-3.308836398	5.19E-08	0.000077
100284943	limonoid UDP-glucosyltransferase	-3.299269237	4.27E-05	0.01371
100283442	glucan endo-1,3-beta-glucosidase A6	-3.265249989	0.000605036	0.088934
100277664	uncharacterized LOC100277664	-3.224030698	0.001553979	0.091948
100275449	uncharacterized LOC100275449	-3.169571427	0.000216026	0.026847
100194409	phosphatidate cytidylyl-transferase	-3.106288015	0.000430627	0.07479
100279414	Protein FLX-like 1	-3.10362114	0.001600825	0.09748
100384583	Protein Chromatin remodeling 24	-3.018440188	0.000170432	0.026508
103630101	LRR receptor-like serine/threonine-protein kinase GSO1	-2.972366806	0.000248077	0.03961
103639870	uncharacterized LOC103639870	-2.913623907	0.000347997	0.050509
100273730	uncharacterized LOC100273730	-2.825396451	0.000693767	0.077925
542019	Gene Zmm16 - Floral homeotic protein PISTILLATA (Uniprot - Q9AR51_MAIZE)	-2.814190945	0.00080591	0.10005
100274431	BTB/POZ domain-containing protein POB1	-2.754769619	6.69E-05	0.018768
100274271	Nicotinate phosphoribosyltransferase (Uniprot - B4FZV8_MAIZE)	-2.50855315	0.000805003	0.059041

101027258	60S ribosomal protein L3	-2.47829604	8.47E-07	0.001599
100501315	uncharacterized LOC100501315	-2.466306173	0.001065043	0.076714
100282596	uncharacterized LOC100282596	-2.442291263	0.000848657	0.077925
100279366	uncharacterized LOC100279366	-2.376020734	0.000475179	0.061366
100381550	uncharacterized LOC100381550	-2.311362235	3.04E-05	0.008515
103650575	histone H3.2	-2.301656929	9.44E-05	0.022124
103632291	phosphomannomutase/phosphoglucomutase	-2.282325714	0.00012397	0.019808
100274030	uncharacterized LOC100274030	-2.194284325	0.000340604	0.03961
109940945	uncharacterized LOC109940945	-2.014946466	0.001812217	0.095106
100283169	Gibberellin 20 oxidase 4 (Uniprot - B6TF99_MAIZE)	1.905836805	0.000258603	0.032714
100286027	uncharacterized LOC100286027	2.079387977	0.000362407	0.051709
100279908	uncharacterized LOC100279908	2.247107978	0.000305175	0.030505
100273789	Villin-2	2.262166446	0.000941003	0.069338
103634514	Abscisic acid receptor PYL4 - like	2.386251179	0.001771926	0.098194
103626511	Protein transport protein SEC16B-like protein	2.413245274	0.001242835	0.099643
100279996	Beta-D-xylosidase 4	2.422343256	0.001684922	0.078197
100384128	putative WRKY transcription factor 34	2.483066281	0.000519928	0.049512
100284121	zinc transporter 4	2.561202902	0.00205315	0.088934
100281089	uncharacterized LOC100281089	2.613041489	6.30E-05	0.02447
103629296	polygalacturonase inhibitor	2.707423171	0.000249085	0.036186
100193995	oligomeric Golgi complex component-related protein	2.711791544	1.56E-05	0.008515
100191528	Late embryogenesis abundant protein group 2	2.768634324	0.000809379	0.072691
100286027	uncharacterized LOC100286027	2.857201666	0.000477097	0.045382
100283092	cell division control protein 50	2.895229428	0.000261273	0.03961
103629667	uncharacterized LOC103629667	2.962789317	0.00074879	0.070052
100280563	acid phosphatase/vanadium-dependent haloperoxidase related	3.068394007	0.001099232	0.076495
100281787	sfp5 - sulfate transporter5	3.120376991	0.001401889	0.10005
103639745	2-oxoglutarate-dependent dioxygenase AOP2	3.14178792	0.000277278	0.088934

	100501347	(chromatin) binding	3.145714903	0.00143684	0.10005
	103646141	kinesin-like protein KIN-7F	3.162697614	0.001488049	0.084377
	100274042	Pectin acetylesterase (Uniprot - B4F9X6_MAIZE)	3.228986405	0.001108907	0.099404
	107546776	uncharacterized LOC107546776	3.294872677	0.000285153	0.032714
	100383384	MACPF domain-containing protein	3.321440368	0.00015843	0.028083
	103625906	glycine-rich cell wall structural protein 1.8	3.35387099	0.000513372	0.064503
	100383398	Calmodulin binding protein	3.414018907	0.00083795	0.077925
	100191163	uncharacterized LOC100191163	3.468039034	0.000417156	0.03961
	100279909	Putative beta-glucosidase 41	3.490406011	0.000693289	0.071575
	100501509	Glycosyltransferase family 61 protein	3.930400521	1.59E-05	0.005618
	100384114	uncharacterized LOC100384114	4.094043211	2.52E-05	0.008784
	100192502	Dormancy-associated protein homolog 3	4.294914272	9.20E-05	0.018829
	100192667	Vacuolar-processing enzyme gamma-isozyme	4.414326535	9.02E-06	0.005438
	542109	uncharacterized LOC542109	4.610843815	9.99E-05	0.020592
	100285036	NAD(P)-binding Rossmann-fold superfamily protein (Uniprot - A0A1D6GD81_MAIZE)	5.070981635	0.000218694	0.098194
	103638313	peroxidase 1	5.1951828	0.00094763	0.090549
	100272285	CRS2-associated factor 1 mitochondrial	20.10703899	2.22E-06	0.010616
	103638687	transcription factor MYB29	20.98768324	2.75E-09	0.000007
GIV vs GII	109944009	uncharacterized LOC109944009	-4.829157595	1.12E-19	1.35E-15
	103654463	UDP-glycosyltransferase 89B2	-3.813570189	1.47E-06	0.008628013
	100274850	uncharacterized LOC100274850	-3.263637464	8.66E-05	0.074589714
	100285166	uncharacterized LOC100285166	-3.204126327	8.77E-05	0.061676621
	100384128	putative WRKY transcription factor 34	2.957987635	3.79E-05	0.047599845
	100192667	Vacuolar-processing enzyme gamma-isozyme	3.762967661	3.88E-05	0.047599845
	109945065	atherin-like	4.574858702	1.36E-05	0.040807546

	100272285	CRS2-associated factor 1 mitochondrial	20.78818688	9.94E-07	0.005957392
GIV vs GIII	100283092	cell division control protein 50	3.136383669	5.60E-05	0.072636
	100384128	putative WRKY transcription factor 34	3.319173225	4.66E-06	0.018467
	103634514	Abscisic acid receptor PYL4 - like	3.338990236	1.63E-05	0.028927
	100280667	uncharacterized LOC100280667	3.386552811	2.50E-05	0.030831
	100384114	uncharacterized LOC100384114	4.10360128	1.70E-05	0.028927
	100282616	(Defective in cullin neddylation protein) DCN1 -like protein 4	4.145947749	3.54E-05	0.053526
	103646141	kinesin-like protein KIN-7F	4.759109263	6.00E-06	0.018467

Table S2. *Azospirillum brasilense* DEGs in each experimental condition.

Experimental condition	Locus tag	Description	log2 Fold-Change	p-value
GIV vs GIII	AMK58_17355	LysR family transcriptional regulator	-3.180995131	0.060720305
	AMK58_14300	glycosyltransferase	2.122879268	0.088023336
	AMK58_16645	amino acid ABC transporter permease	2.808972249	0.088782658
	AMK58_01595	hypothetical protein	3.36040899	0.08873343
	AMK58_07645	amino acid ABC transporter substrate-binding protein	3.510425081	0.0695893
	AMK58_11375	type I methionyl aminopeptidase	4.114762172	0.076177057
	AMK58_18290	ketoacyl-ACP synthase III	4.413982956	0.09364138
	AMK58_20205	FUSC family protein	4.418102723	0.064819336
	AMK58_21350	DUF882 domain-containing protein	4.430487027	0.087962317
	AMK58_03600	PrkA family serine protein kinase	4.462374506	0.025934774
	AMK58_27405	sugar ABC transporter permease	4.482661646	0.078481562
	AMK58_20090	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase	5.418773104	0.005594081
	AMK58_14270	iron ABC transporter substrate-binding protein	6.483718319	0.097448745
	AMK58_06945	oxalate/formate MFS antiporter	6.686216682	0.000443133

Table S3. Maize's differential expressed uncharacterized genes in each experimental condition.

Experimental Condition	Gene ID	Description	log2 Fold-Change	p-value	p-adj
GII vs GI	103646825	uncharacterized LOC103646825	-5.971302807	1.53E-05	0.038420757
	103644253	uncharacterized LOC103644253	-5.458778999	4.28E-05	0.038420757
	111590293	uncharacterized LOC111590293	-4.61713408	1.74E-05	0.038420757
	100280286	uncharacterized LOC100280286	-4.448925606	9.46E-06	0.02229954
	100275449	uncharacterized LOC100275449	-3.541209705	4.33E-05	0.037875323
	103655263	uncharacterized LOC103655263	-3.023127338	0.00015546	0.087565331
	100381550	uncharacterized LOC100381550	-2.277649056	4.13E-05	0.032770133
GIII vs GI	103644253	uncharacterized LOC103644253	-8.771883724	1.05E-07	0.000340361
	103646825	uncharacterized LOC103646825	-6.381008475	4.59E-06	0.040656599
	100279341	uncharacterized LOC100279341	-5.44058078	0.000221553	0.053829856
	109944009	uncharacterized LOC109944009	-5.347180688	2.05E-25	1.30E-21
	100280286	uncharacterized LOC100280286	-4.119236592	3.01E-05	0.013078967
	100274041	uncharacterized LOC100274041	-4.008946848	0.000325405	0.077984566
	100277664	uncharacterized LOC100277664	-3.99506072	0.00013732	0.031031202
	100192499	uncharacterized LOC100192499	-3.982224712	0.000226684	0.067665291
	100275449	uncharacterized LOC100275449	-2.966525509	0.000523379	0.08166411
	100381550	uncharacterized LOC100381550	-2.369286588	1.96E-05	0.011025741
	100502224	uncharacterized LOC100502224	-2.131075988	1.36E-05	0.011332009
	100382848	uncharacterized LOC100382848	3.969145051	0.000180479	0.053829856
GIV vs GI	103644253	uncharacterized LOC103644253	-7.840435947	2.00E-06	0.001902
	100384061	uncharacterized LOC100384061	-6.673609541	0.000854958	0.07479
	109944009	uncharacterized LOC109944009	-6.27415482	3.43E-32	4.52E-28
	103646825	uncharacterized LOC103646825	-5.592077697	7.33E-06	0.010616
	100274041	uncharacterized LOC100274041	-4.719772344	4.20E-05	0.010616
	100192499	uncharacterized LOC100192499	-4.451092579	5.05E-05	0.018829
	111590293	uncharacterized LOC111590293	-4.3444928	2.07E-05	0.034623
	103653702	uncharacterized LOC103653702	-3.979243115	0.000209593	0.060634
	100193544	uncharacterized LOC100193544	-3.523498625	0.000138305	0.028083
	100280286	uncharacterized LOC100280286	-3.471764588	0.00035198	0.032714
	100381815	uncharacterized LOC100381815	-3.448432964	0.000279114	0.046466
	100383495	uncharacterized LOC100383495	-3.374979564	0.000169831	0.035184
	100277664	uncharacterized LOC100277664	-3.224030698	0.001553979	0.091948
	100275449	uncharacterized LOC100275449	-3.169571427	0.000216026	0.026847
	103639870	uncharacterized LOC103639870	-2.913623907	0.000347997	0.050509
	100273730	uncharacterized LOC100273730	-2.825396451	0.000693767	0.077925
	100501315	uncharacterized LOC100501315	-2.466306173	0.001065043	0.076714

	100282596	uncharacterized LOC100282596	-2.442291263	0.000848657	0.077925
	100279366	uncharacterized LOC100279366	-2.376020734	0.000475179	0.061366
	100381550	uncharacterized LOC100381550	-2.311362235	3.04E-05	0.008515
	100274030	uncharacterized LOC100274030	-2.194284325	0.000340604	0.03961
	109940945	uncharacterized LOC109940945	-2.014946466	0.001812217	0.095106
	100286027	uncharacterized LOC100286027	2.079387977	0.000362407	0.051709
	100279908	uncharacterized LOC100279908	2.247107978	0.000305175	0.030505
	100281089	uncharacterized LOC100281089	2.613041489	6.30E-05	0.02447
	100286027	uncharacterized LOC100286027	2.857201666	0.000477097	0.045382
	103629667	uncharacterized LOC103629667	2.962789317	0.00074879	0.070052
	107546776	uncharacterized LOC107546776	3.294872677	0.000285153	0.032714
	100191163	uncharacterized LOC100191163	3.468039034	0.000417156	0.03961
	100384114	uncharacterized LOC100384114	4.094043211	2.52E-05	0.008784
	542109	uncharacterized LOC542109	4.610843815	9.99E-05	0.020592
GIV vs GII	109944009	uncharacterized LOC109944009	-4.829157595	1.12E-19	1.35E-15
	100274850	uncharacterized LOC100274850	-3.263637464	8.66E-05	0.074589714
	100285166	uncharacterized LOC100285166	-3.204126327	8.77E-05	0.061676621
GIV vs GIII	100280667	uncharacterized LOC100280667	3.386552811	2.50E-05	0.030831
	100384114	uncharacterized LOC100384114	4.10360128	1.70E-05	0.028927

5. Análise Complementar

Ontologia Gênica e Enriquecimento de Rotas Metabólicas dos genes diferencialmente expressos (DEGs)

Com a finalidade de atribuir o papel das proteínas codificadas pelos genes diferencialmente expressos identificados, tanto para milho como para *A. brasilense*, foram realizadas análises de grupos ontológicos (GOs) e de enriquecimento de rotas metabólicas.

Para se realizar o teste de enriquecimento para perfis funcionais de categorias ontológicas gênicas e o teste de enriquecimento das vias metabólicas usando o banco de dados KEGG (*Kyoto Encyclopedia of Genes and Genomes*, em inglês), foi utilizado o pacote clusterProfiler do programa R. Esses pacote realiza esses testes baseado na distribuição hipergeométrica dos DEGs identificados em ambos os organismos (p-valor < 0.05 e q < 0.2) (Yu *et al.* 2012; Li *et al.* 2019). Para se realizar os testes dos perfis ontológicos, as informações das anotações dos genomas de *Z. mays* e *A. brasilense* foram baixadas usando o pacote AnnotationHub do programa R (http://bioconductor.org/packages/release/bioc/html/AnnotationHub.html).

6. Discussão

O milho é uma das culturas mais importantes para alimentação humana e animal. Segundo a *Food and Agriculture Organization of the United Nations* (FAO, 2016), em 2016, o Brasil apresentou a 3ª maior produção mundial. Dada a importância nutricional e econômica desta cultura, pesquisar formas de aumentar sua produção se faz necessária. Uma alternativa para se aumentar a produção desta cultura sem aumento na adição de nitrogênio é o uso de bactérias promotoras de crescimento vegetal (PGPB). Essas bactérias são capazes de estimular um maior crescimento vegetal por meio da fixação biológica de nitrogênio (FBN), da produção de fitormônios, vitaminas e fatores de crescimento, que estimulam o crescimento e a produção das plantas (Babalola, 2010; Hungria *et al.* 2010; Bashan and de-Bashan, 2010).

Apesar de já estar bem estabelecido o benefício de *Azospirillum* para as plantas, pouco é sabido em relação à expressão gênica dos organismos durante a interação. Para lançar um pouco de luz nessa relação e como o AIA produzido pela bactéria influencia a planta, estudamos as alterações no transcritoma da planta e da bactéria durante na presença de Yucasin, que é um potente inibidor competitivo na planta da conversão de AIP em AIA pela via TAA/YUC.

O primeiro desafio desse trabalho foi separar os transcritomas dos dois organismos que estavam interagindo, uma vez que as amostras de RNA continham materiais tanto do milho como da bactéria. A literatura recomenda que essa separação seja feita *in silico*, durante a fase de mapeamento dos mRNA (Wolf *et al.* 2018). Para tanto, há duas metodologias que podem ser escolhidas: sequencial e combinada. Apesar de Wolf *et al.* (2018) recomendarem a segunda alternativa, por permitir que o programa de mapeamento defina em qual dos genomas a sequência irá se alinhar de forma mais eficiente, a análise sequencial ainda é a mais usada na maioria dos trabalhos de *Dual RNA-seq* (Kovalchuk *et al.* 2019; LaMonte *et al.* 2019; Mateus *et al.* 2019; Montoya *et al.* 2019; Mutha *et al.* 2019). Acreditamos que isso ocorra devido ao fato de que, para se criar uma referência concatenada com os genomas dos organismos presentes na amostra, para que os programas de mapeamento mais utilizados possam usá-la no mapeamento, é necessário um conhecimento mais apurado de programação. Aprianto *et al.* (2016), por exemplo, adicionaram o genoma de *Pseudomonas* como um cromossomo extra no genoma de referência de *Homo sapiens*, fato que exigiu um conhecimento mais aprofundado de bioinformática. Para contornar esse problema, desenvolvemos uma abordagem de mapeamento combinada, usando o programa CLC Genomics Workbench, utilizando bibliotecas de RNA disponíveis em bancos de dados públicos e de um experimento conduzido no nosso laboratório. Graças às características do programa, realizamos a separação *in silico* durante a etapa de mapeamento dos *reads* obtidos. Durante essa etapa foi usado um arquivo contendo os dois genomas de referência, que foram previamente

concatenados usando o terminal do Linux. Depois de mapeados, o programa nos permitiu extrair as sequências que mapearam em cada genoma (Capítulo 1, Figura S3). Feito isso, cada biblioteca pode ter suas sequências contadas usando o genoma de referência com as respectivas anotações.

Nossos resultados mostraram que a metodologia combinada é, de fato, a melhor opção para análises de dados de *Dual RNA-seq*, visto que devido ao uso dela houve a diminuição do número de reads que mapearam cruzado. Além disso essa abordagem oferece uma alternativa para a realização deste procedimento em um ambiente amigável ao usuário e sem a necessidade de grandes habilidades na área da programação.

Uma vez definida a forma de análise, foram realizados os isolamentos dos RNAs totais e as bibliotecas de cDNA obtidas foram sequenciadas. As sequências foram mapeadas contra os respectivos genomas de referência e os *loci* em que essas sequências mapearam foram determinados. Por fim, foram identificados, entre os *loci* que tiveram sequências atribuídas a eles, os genes diferencialmente expressos entre os grupos experimentais. Dessa análise foram identificados genes diferencialmente expressos no milho que codificam proteínas envolvidas na resposta a estresses bióticos e abióticos e envolvidas na síntese e resposta a fito-hormônios.

Da análise dos genes envolvidos na resposta e síntese de fito-hormônios foram identificados genes pertencentes às vias do ácido abscísico (ABA) e das giberelinas (GA). Com relação à via de resposta ao ABA, foram identificados três genes diferencialmente expressos. Com base na análise da expressão desses genes, elaboramos uma proposta de como a presença de *A. brasilense* na planta melhora a resposta dessa ao estresse abiótico (Capítulo 2, Figura 4). Também foi identificado um gene que codifica o receptor de GA *GID1L2* (GIBBERELLIN-INSENSITIVE DWARF1). Esse receptor, quando está ligado à GA, se liga à proteína DELLA, marcando-a para degradação pelo proteassoma 26S (Hirano *et al.* 2008). A expressão do gene *GID1L2* em ambas as situações em que ele foi identificado (grupos II e IV em relação ao grupo I) estava reprimida. No grupo IV, onde também havia a presença de *Azospirillum*, a repressão foi uma ordem de grandeza menor. Na comparação do grupo IV com o grupo I também foi identificado que o gene que codifica a *Gibberellin 20-oxidase 4* (GA20oxi4), que, apesar de ter um valor válido de p-adjustado, esse gene apresentou um valor de expressão muito próximo do limite estabelecido por nós, ficando em 1,9. Com base nas expressões observadas para esses dois genes e pelo fato de que *A. brasilense* sintetiza AIA durante a interação com a planta (Duca *et al.* 2014), podemos sugerir que a auxina produzida pela bactéria de alguma forma estimulou a síntese de *GID1L2*, visto que na comparação do grupo que somente recebeu Yucasin (grupo II), essa repressão foi mais forte do que no grupo que recebeu Yucasin na presença da bactéria (grupo IV). Com relação ao observado para o gene *ZmGA20ox4*,

levando em consideração que a presença da bactéria e de Yucasin estimulou a síntese de ABA (Capítulo 2, figura 4), podemos supor que, com base no que postula Hirano *et al.* (2008), o ABA produzido interagiu com DELLA, estabilizando-a, e, assim, permitiu a expressão de *ZmGA20ox4*. Estes dois genes não foram incluídos na discussão do artigo no segundo capítulo devido ao fato de que novas análises são necessárias para confirmar nossas observações para *ZmGA20ox4*.

Com relação aos genes envolvidos na resposta a estresses bióticos e abióticos, vale destacar a expressão dos genes *TYFY 3B* e *WIN1*. Ambos, quando analisamos o grupo que somente recebeu a bactéria (III) contra o grupo controle (I), estavam reprimidos. Isso indica que, de alguma forma, a presença de *A. brasilense* reprimiu a expressão desses genes. Como eles estão relacionados com a resposta a estresses bióticos podemos supor que sua repressão facilita a colonização da planta pela bactéria.

Em nossas análises também foi identificado o gene que codifica o receptor *GASSHO 1* (*GSO1*). Esse receptor é um tipo de quinase receptora rica em resíduos de leucina (*leucine-rich repeat receptor-like kinase* – LRR-RLK). Esse gene foi identificado em mutantes de *Arabidopsis* e atua junto com *GSO2* (Tsuwamoto *et al.* 2008). Um duplo mutante para esses genes apresentou diversos danos na raiz, entre eles, o aumento da permeabilidade da epiderme das raízes (Racolta *et al.* 2014; Tsuwamoto *et al.* 2008). Como não foram observadas alterações na expressão de *GSO2* e como *GSO1* estava reprimido somente nos grupos que em que havia a presença da bactéria, postulamos que esse receptor possa estar envolvido na facilitação das trocas de substâncias entre a planta e a bactéria. Esse envolvimento pode se dar através do aumento da permeabilidade celular sem, contudo, haver dano ao crescimento das raízes, como observado por Racolta *et al.* (2014), visto que o grupo III apresentou tamanho de raiz compatível com os tamanhos de raízes das plantas dos grupos I e IV. As raízes das plantas do grupo IV, inclusive, apresentaram crescimento maior que aquelas dos grupos I e II (Capítulo 2, Figura 2A).

Por fim, vale destacar que, graças à abordagem de análise conjunta desenvolvida nesse trabalho, foi possível a detecção de genes diferencialmente expressos em *A. brasilense*. Apesar da quantidade de genes detectados ter sido pequena, os genes identificados indicaram que a bactéria estava ativa. Essa pequena quantidade de genes detectados provavelmente ocorreu porque antes de se preparar as bibliotecas de cDNA para o sequenciamento foi usado um *kit* para a remoção de rRNA exclusivo para plantas (RiboMinus™ Plant Kit for RNA-Seq - Invitrogen). Muito provavelmente esta estratégia não foi suficiente para eliminar o rRNA bacteriano. Levantamos essa hipótese porque, segundo Petrova *et al.* (2017), a quantidade de rRNA é muito maior que a de mRNA nas amostras bacterianas e a metodologia de remoção utilizada pode alterar a qualidade

dessa remoção e, consequentemente, afetar as subsequentes análises. Isso foi observado em nossas amostras na medida em que a maior parte das sequências bacterianas alinharam em regiões correspondentes aos genes codificadores de rRNA. Assim, para futuros estudos de *Dual RNA-seq* que envolvam a interação eucarionte – procarionte recomenda-se que, antes de se preparar o cDNA, sejam realizadas duas etapas de depleção para remover tanto rRNA procarionte quanto eucarionte, a fim de aumentar a eficiência no sequenciamento do mRNA bacteriano.

7. Conclusões Finais e Perspectivas

Na primeira parte deste trabalho foi aprimorada uma abordagem de análise que foi posteriormente utilizada para analisar as bibliotecas de cDNA obtidas. Dessa parte do trabalho foi possível concluir que em experimentos de *Dual RNA-seq*, que são formados por sequências de RNA de dois ou mais organismos, é preferível o uso da metodologia combinada para analisar esses dados. A metodologia sequencial pode levar a erros de sub e/ou super-estimação da expressão de genes em cada um dos organismos a depender de qual deles foi usado primeiro no mapeamento. Outro resultado obtido foi mostrar uma forma de executar esta metodologia em um ambiente amigável para o usuário e que demanda pouco conhecimento de informática para ser realizada.

Na segunda parte do trabalho foi estudada a interação entre *Z. mays* e *A. brasilense* FP2 durante a inibição da produção de AIA pela planta causada pelo composto Yucasin. Com base nos dados fisiológicos obtidos foi possível observar que os efeitos benéficos de *A. brasilense* FP2 foram capazes de reverter o fenótipo das raízes de milho causado pelo Yucasin (Capítulo 2, Figura S2-D). Da análise dos dados de expressão gênica foi possível concluir que a presença de *A brasilense* provavelmente melhorou a resposta da planta a estresses abióticos, através da via de resposta dependente de ABA. Também foi observado, na presença somente da bactéria, que os genes que codificam o fator de transcrição *TYF 3B* e a proteína *WIN1,* envolvidos na resposta a estresse biótico, estavam com sua expressão reprimida. Concluímos que, de alguma forma, a presença da bactéria causou esta repressão, o que, provavelmente, facilitou a colonização da planta pela bactéria.

Por fim, identificamos que um dos genes da via de síntese de GA estava com sua expressão estimulada, mas abaixo dos parâmetros usados no nosso experimento. Entretanto, esse valor estava somente 0,1 ponto abaixo do limite, o que nos leva a crer que esse estímulo poderia corresponder à realidade. Também foram identificados em nossas análises um total de 55 genes não caracterizados em todas as situações experimentais.

Como tivemos mais de 170 genes diferencialmente expressos entre genes de milho e de *Azospirillum* (Capítulo 2, Tabelas S1 e S2), análises complementares desses genes já estão sendo realizadas. Foram realizados até o presente momento o teste de enriquecimento para perfis funcionais de categorias ontológicas gênicas (GO) e o teste de enriquecimento das vias metabólicas usando o banco de dados KEGG (Tabelas Suplementares 1 e 2). Análises preliminares indicam que parte deles mapearam em rotas metabólicas e em categorias ontológicas, especialmente na de processos biológicos (Tabela Suplementar 2).

As próximas etapas envolverão a realização de análises complementares por PCR em tempo real do gene *Zm20ox4* e de outros da via das GA para comprovar os resultados apresentados.

Também serão finalizadas as análises dos testes de GO e de rotas metabólicas. Finalmente, alguns dos genes não caracterizados serão investigados. Será dada especial atenção nos DEGs não caracterizados identificados na comparação do grupo com a bactéria (III) com o grupo I. Com exceção de um, todos os demais DEGs não caracterizados estavam reprimidos nessa comparação.

8. Tabelas suplementares

Table S1. KEGG enrichment pathways identified in each experimental condition to maize and *Azospirillum brasilense*.

Experimental Condition	Pathway ID	Description	Gene Ratio	BgRatio	p-value	p-adjust	q-value	Gene ID
Zea mays								
GIII vs GI	zma04070	Phosphatidyl-inositol signaling system	2\11	95\6705	0.010053	0.15080161	0.137573399	100194409\ 103643201
GIV vs GI	zma00940	Phenylpropanoid biosynthesis	4\16	243\6708	0.002167	0.04767758	0.043343255	100383495\ 100279909\ 100501315\ 103638313
	zma00460	Cyanoamino acid metabolism	2\16	61\6708	0.008993	0.09892668	0.089933348	100279909\ 100501315
	zma00630	Glyoxylate and dicarboxylate metabolism	2\16	101\6708	0.02348	0.17218497	0.156531795	103636394\ 100281089
GIV vs GII	zma04626	Plant-pathogen interaction	2\2	209\6708	0.000966	0.00096625	NA	100384128\ 100285166
Azospirillum brasilense								
Group IV vs Group III	abf02010	ABC transporters	3\5	212\1789	0.01368	0.08210	0.05761	AMK58_07645\ AMK58_27405\ AMK58_14270
	abf00900	Terpenoid backbone biosynthesis	1\5	11\1789	0.03040	0.09120	0.06400	AMK58_20090

Table S2. GO enrichment categories identified in each experimental condition to maize or *Azospirillum brasilense.*

Experimental Condition	GO Category	Pathway ID	Description	Gene Ratio	p-value	p-adjust	q-value	Gene ID
Zea mays								
GII vs GI	Biological Process	GO:0001522	pseudouridine synthesis	1\4	0.010743	0.08431	0.041909	100217137
		GO:0042742	defense response to bacterium	1\4	0.011762	0.08431	0.041909	103652813
		GO:0009617	response to bacterium	1\4	0.013796	0.08431	0.041909	103652813
		GO:0050832	defense response to fungus	1\4	0.014305	0.08431	0.041909	103652813
		GO:0009620	response to fungus	1\4	0.016843	0.08431	0.041909	103652813
		GO:0006220	pyrimidine nucleotide metabolic process	1\4	0.021399	0.08431	0.041909	100194409
		GO:0006221	pyrimidine nucleotide biosynthetic process	1\4	0.021399	0.08431	0.041909	100194409
		GO:0046474	glycerophospholipid biosynthetic process	1\4	0.022914	0.08431	0.041909	100194409
		GO:0045017	glycerolipid biosynthetic process	1\4	0.024428	0.08431	0.041909	100194409
		GO:0072528	pyrimidine-containing compound biosynthetic process	1\4	0.028958	0.08431	0.041909	100194409
		GO:0072527	pyrimidine-containing compound metabolic process	1\4	0.029962	0.08431	0.041909	100194409
		GO:0008654	phospholipid biosynthetic process	1\4	0.032971	0.08431	0.041909	100194409

		GO:0098542	defense response to other organism	1\4	0.032971	0.08431	0.041909	103652813
		GO:0009451	RNA modification	1\4	0.037471	0.08431	0.041909	100217137
		GO:0043207	response to external biotic stimulus	1\4	0.037471	0.08431	0.041909	103652813
		GO:0051707	response to other organism	1\4	0.037471	0.08431	0.041909	103652813
		GO:0006650	glycerophospholipid metabolic process	1\4	0.040463	0.084906	0.042205	100194409
		GO:0046486	glycerolipid metabolic process	1\4	0.042453	0.084906	0.042205	100194409
		GO:0009607	response to biotic stimulus	1\4	0.045433	0.086083	0.04279	103652813
	Molecular Function	GO:0009982	pseudouridine synthase activity	1\10	0.012964	0.155569	0.136464	100217137
		GO:0016866	intramolecular transferase activity	1\10	0.031754	0.190525	0.167127	100217137
	Cellular Component	GO:0005730	nucleolus	1\3	0.023359	0.096374	0.084538	100283525
GIII vs GI	Biological Process	GO:0006783	heme biosynthetic process	1\9	0.016069	0.166101	0.134966	100192499
		GO:0042168	heme metabolic process	1\9	0.019482	0.166101	0.134966	100192499
		GO:0001522	pseudouridine synthesis	1\9	0.024017	0.166101	0.134966	100217137
		GO:0042742	defense response to bacterium	1\9	0.026278	0.166101	0.134966	103652813

	GO:0009617	response to bacterium	1\9	0.030784	0.166101	0.134966	103652813
	GO:0050832	defense response to fungus	1\9	0.031908	0.166101	0.134966	103652813
	GO:0006779	porphyrin-containing compound biosynthetic process	1\9	0.035273	0.166101	0.134966	100192499
	GO:0009620	response to fungus	1\9	0.03751	0.166101	0.134966	103652813
	GO:0033014	tetrapyrrole biosynthetic process	1\9	0.04197	0.166101	0.134966	100192499
	GO:0006778	porphyrin-containing compound metabolic process	1\9	0.044194	0.166101	0.134966	100192499
	GO:0006220	pyrimidine nucleotide metabolic process	1\9	0.04752	0.166101	0.134966	100194409
	GO:0006221	pyrimidine nucleotide biosynthetic process	1\9	0.04752	0.166101	0.134966	100194409
	GO:0033013	tetrapyrrole metabolic process	1\9	0.049732	0.166101	0.134966	100192499
Molecular Function	GO:0015399	primary active transmembrane transporter activity	2\18	0.009032	0.126454	0.099832	100286099\ 103641772
	GO:0015405	P-P-bond-hydrolysis-driven transmembrane transporter activity	2\18	0.009032	0.126454	0.099832	100286099\ 103641772
	GO:0009678	hydrogen-translocating pyrophosphatase activity	1\18	0.013722	0.128071	0.101109	100286099
	GO:0009982	pseudouridine synthase activity	1\18	0.023221	0.137599	0.108631	100217137
	GO:0004427	inorganic diphosphatase activity	1\18	0.024571	0.137599	0.108631	100286099

		GO:0022804	active transmembrane transporter activity	2\18	0.032179	0.150167	0.118553	100286099\ 103641772
		GO:0035091	phosphatidylinositol binding	1\18	0.039305	0.15722	0.124121	100502224
GIV vs GI	Biological Process	GO:0046474	glycerophospholipid biosynthetic process	2\23	0.007645	0.156418	0.117756	100381815\ 100194409
		GO:0045017	glycerolipid biosynthetic process	2\23	0.008664	0.156418	0.117756	100381815\ 100194409
		GO:0008654	phospholipid biosynthetic process	2\23	0.015498	0.156418	0.117756	100381815\ 100194409
		GO:0022402	cell cycle process	2\23	0.02239	0.156418	0.117756	100191163\ 103653702
		GO:0006650	glycerophospholipid metabolic process	2\23	0.022924	0.156418	0.117756	100381815\ 100194409
		GO:0046486	glycerolipid metabolic process	2\23	0.025109	0.156418	0.117756	100381815\ 100194409
		GO:0006900	vesicle budding from membrane	1\23	0.029156	0.156418	0.117756	103626511
		GO:0007093	mitotic cell cycle checkpoint	1\23	0.029156	0.156418	0.117756	103653702
		GO:0045930	negative regulation of mitotic cell cycle	1\23	0.029156	0.156418	0.117756	103653702
		GO:0048194	Golgi vesicle budding	1\23	0.029156	0.156418	0.117756	103626511
		GO:0051784	negative regulation of nuclear division	1\23	0.029156	0.156418	0.117756	103653702
		GO:0090114	COPII-coated vesicle budding	1\23	0.029156	0.156418	0.117756	103626511
		GO:0090407	organophosphate biosynthetic process	3\23	0.031439	0.156418	0.117756	100381815\ 100274271\ 100194409
		GO:0007088	regulation of mitotic nuclear division	1\23	0.032027	0.156418	0.117756	103653702

GO:0007091	metaphase\anaphase transition of mitotic cell cycle	1\23	0.032027	0.156418	0.117756	103653702
GO:0010965	regulation of mitotic sister chromatid separation	1\23	0.032027	0.156418	0.117756	103653702
GO:0019674	NAD metabolic process	1\23	0.032027	0.156418	0.117756	100274271
GO:0030071	regulation of mitotic metaphase\anaphase transition	1\23	0.032027	0.156418	0.117756	103653702
GO:0033045	regulation of sister chromatid segregation	1\23	0.032027	0.156418	0.117756	103653702
GO:0033047	regulation of mitotic sister chromatid segregation	1\23	0.032027	0.156418	0.117756	103653702
GO:0044784	metaphase\anaphase transition of cell cycle	1\23	0.032027	0.156418	0.117756	103653702
GO:0051306	mitotic sister chromatid separation	1\23	0.032027	0.156418	0.117756	103653702
GO:0051983	regulation of chromosome segregation	1\23	0.032027	0.156418	0.117756	103653702
GO:1902099	regulation of metaphase\anaphase transition of cell cycle	1\23	0.032027	0.156418	0.117756	103653702
GO:1905818	regulation of chromosome separation	1\23	0.032027	0.156418	0.117756	103653702
GO:0007010	cytoskeleton organization	2\23	0.033394	0.156418	0.117756	100191163\ 100273789
GO:0006996	organelle organization	4\23	0.033663	0.156418	0.117756	100191163\ 103653702\ 100273789\ 103626511
GO:0010948	negative regulation of cell cycle process	1\23	0.034889	0.156418	0.117756	103653702
GO:0016050	vesicle organization	1\23	0.034889	0.156418	0.117756	103626511

	GO:0044770	cell cycle phase transition	1\23	0.034889	0.156418	0.117756	103653702
	GO:0044772	mitotic cell cycle phase transition	1\23	0.034889	0.156418	0.117756	103653702
	GO:1901987	regulation of cell cycle phase transition	1\23	0.034889	0.156418	0.117756	103653702
	GO:1901990	regulation of mitotic cell cycle phase transition	1\23	0.034889	0.156418	0.117756	103653702
	GO:2001251	negative regulation of chromosome organization	1\23	0.034889	0.156418	0.117756	103653702
	GO:0000075	cell cycle checkpoint	1\23	0.037743	0.156418	0.117756	103653702
	GO:0061982	meiosis I cell cycle process	1\23	0.037743	0.156418	0.117756	100191163
	GO:0006644	phospholipid metabolic process	2\23	0.038547	0.156418	0.117756	100381815\ 100194409
	GO:0006783	heme biosynthetic process	1\23	0.040589	0.156418	0.117756	100192499
	GO:0007051	spindle organization	1\23	0.040589	0.156418	0.117756	100191163
	GO:0051304	chromosome separation	1\23	0.040589	0.156418	0.117756	103653702
	GO:0051656	establishment of organelle localization	1\23	0.040589	0.156418	0.117756	103626511
	GO:0007017	microtubule-based process	2\23	0.042601	0.160261	0.12065	100191163\ 103646141
	GO:0051783	regulation of nuclear division	1\23	0.046258	0.16997	0.127959	103653702
	GO:0042168	heme metabolic process	1\23	0.04908	0.174493	0.131364	100192499
	GO:0007049	cell cycle	2\23	0.049697	0.174493	0.131364	100191163\ 103653702
Molecular Function	GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	4\33	0.008857	0.176761	0.148851	100279909\ 100501315\ 100279996\ 100283442

		GO:0016798	hydrolase activity, acting on glycosyl bonds	4\33	0.011962	0.176761	0.148851	100279909\ 100501315\ 100279996\ 100283442
		GO:0005385	zinc ion transmembrane transporter activity	1\33	0.025027	0.176761	0.148851	100284121
		GO:0009678	hydrogen-translocating pyrophosphatase activity	1\33	0.025027	0.176761	0.148851	100286099
		GO:0072509	divalent inorganic cation transmembrane transporter activity	1\33	0.025027	0.176761	0.148851	100284121
		GO:1901682	sulfur compound transmembrane transporter activity	1\33	0.025027	0.176761	0.148851	100281787
		GO:0051015	actin filament binding	1\33	0.027496	0.176761	0.148851	100273789
		GO:0046915	transition metal ion transmembrane transporter activity	1\33	0.039751	0.223053	0.187834	100284121
		GO:0004427	inorganic diphosphatase activity	1\33	0.044611	0.223053	0.187834	100286099
GIV vs GIII	Molecular Function	GO:0001871	pattern binding	1\3	0.009631	0.031244	0.006578	100280667
		GO:0030247	polysaccharide binding	1\3	0.009631	0.031244	0.006578	100280667
		GO:0003777	microtubule motor activity	1\3	0.010089	0.031244	0.006578	103646141
		GO:0003774	motor activity	1\3	0.013512	0.031244	0.006578	103646141
		GO:0008017	microtubule binding	1\3	0.016701	0.031244	0.006578	103646141
		GO:0015631	tubulin binding	1\3	0.018747	0.031244	0.006578	103646141
		GO:0008092	cytoskeletal protein binding	1\3	0.035237	0.050339	0.010598	103646141
		GO:0030246	carbohydrate binding	1\3	0.042854	0.053568	0.011277	100280667

	Cellular Component	GO:0005874	microtubule	1\5	0.039673	0.061203	0.016106	103646141
		GO:0099080	supramolecular complex	1\5	0.040194	0.061203	0.016106	103646141
		GO:0099081	supramolecular polymer	1\5	0.040194	0.061203	0.016106	103646141
		GO:0099512	supramolecular fiber	1\5	0.040194	0.061203	0.016106	103646141
		GO:0099513	polymeric cytoskeletal fiber	1\5	0.040194	0.061203	0.016106	103646141
		GO:0015630	microtubule cytoskeleton	1\5	0.045902	0.061203	0.016106	103646141
Azospirillum brasilense								
Group IV vs Group III	Biological Process	GO:0008610	lipid biosynthetic process	2\5	0.005961	0.147531	0.13059	AMK58_RS20105\ AMK58_RS18310
		GO:0044255	cellular lipid metabolic process	2\5	0.00748	0.147531	0.13059	AMK58_RS20105\ AMK58_RS18310
		GO:0006629	lipid metabolic process	2\5	0.010059	0.147531	0.13059	AMK58_RS20105\ AMK58_RS18310
		GO:0032787	monocarboxylic acid metabolic process	2\5	0.01588	0.174676	0.154618	AMK58_RS20105\ AMK58_RS18310
		GO:0006720	isoprenoid metabolic process	1\5	0.028074	0.205874	0.182233	AMK58_RS20105
		GO:0008299	isoprenoid biosynthetic process	1\5	0.028074	0.205874	0.182233	AMK58_RS20105
		GO:0006633	fatty acid biosynthetic process	1\5	0.045608	0.230591	0.204111	AMK58_RS18310

9. Referências Bibliográficas

Aprianto, R., Slager, J., Holsappel, S. and Veening, J.-W. (2016) Time-resolved dual RNA-seq reveals extensive rewiring of lung epithelial and pneumococcal transcriptomes during early infection, Genome Biol, 17(1), pp. 198.

Arcondeguy, T., Jack, R. and Merrick, M. (2001) P(II) signal transduction proteins, pivotal players in microbial nitrogen control, Microbiol Mol Biol Rev, 65(1), pp. 80-105.

Babalola, O. O. (2010) Beneficial bacteria of agricultural importance, Biotechnol Lett, 32(11), pp. 1559-70.

Baddal, B., Muzzi, A., Censini, S., Calogero, R. A., Torricelli, G., Guidotti, S., Taddei, A. R., Covacci, A., Pizza, M., Rappuoli, R., Soriani, M. and Pezzicoli, A. (2015) Dual RNA-seq of Non type able *Haemophilus influenzae* and Host Cell Transcriptomes Reveals Novel Insights into Host-Pathogen Cross Talk, mBio, 6(6), pp. e01765-15.

Balsanelli, E., Tadra-Sfeir Michelle, Z., Faoro, H., Pankievicz Vânia, C. S., Baura Valter, A., Pedrosa Fábio, O., Souza Emanuel, M., Dixon, R. and Monteiro Rose, A. (2016) Molecular adaptations of *Herbaspirillum seropedicae* during colonization of the maize rhizosphere, Environ Microbiol, 18(8), pp. 2343-2356.

Barret, M., Frey-Klett, P., Guillerm-Erckelboudt, A.-Y., Boutin, M., Guernec, G. and Sarniguet, A. (2009) Effect of Wheat Roots Infected with the Pathogenic Fungus *Gaeumannomyces graminis* var. tritici on Gene Expression of the Biocontrol Bacterium *Pseudomonas* fluorescens Pf29Arp, Mol Plant Microbe In, 22(12), pp. 1611-1623.

Bashan, Y. and de-Bashan, L. E. (2010) How the Plant Growth-Promoting Bacterium *Azospirillum* Promotes Plant Growth—A Critical Assessment, 108, pp. 77-136.

Baudoin, E. L., A.;Mirza, M. S.;El Zemrany, H.;Prigent-Combaret, C.;Jurkevich, E.;Spaepen, S.;Vanderleyden, J.;Nazaret, S.;Okon, Y.;Moenne-Loccoz, Y. (2010) Effects of *Azospirillum brasilense* with genetically modified auxin biosynthesis gene *ipdC* upon the diversity of the indigenous microbiota of the wheat rhizosphere, Res Microbiol, 161(3), pp. 219-26.

Bishopp, A., Mahonen, A. P. and Helariutta, Y. (2006) Signs of change: hormone receptors that regulate plant development, Development, 133(10), pp. 1857-69.

Blakeslee, J. J., Peer, W. A. and Murphy, A. S. (2005) Auxin transport, Curr Opin Plant Biol, 8(5), pp. 494-500.

Bonato, P., Batista Marcelo, B., Camilios-Neto, D., Pankievicz Vânia, C. S., Tadra-Sfeir Michelle, Z., Monteiro Rose, A., Pedrosa Fabio, O., Souza Emanuel, M., Chubatsu Leda, S., Wassem, R.

and Rigo Liu, U. (2016) RNA-seq analyses reveal insights into the function of respiratory nitrate reductase of the diazotroph *Herbaspirillum seropedicae*, Environ Microbiol, 18(8), pp. 2677-2688.

Boscari, A., del Giudice, J., Ferrarini, A., Venturini, L., Zaffini, A.-L., Delledonne, M. and Puppo, A. (2012) Expression dynamics of the *Medicago truncatula* transcriptome during the symbiotic interaction with *Sinorhizobium meliloti*: which role for nitric oxide?, Plant Physiol, 161(1), pp. 425-39.

Bottini, R., Cassan, F. and Piccoli, P. (2004) Gibberellin production by bacteria and its involvement in plant growth promotion and yield increase, Appl Microbiol Biotechnol, 65(5), pp. 497-503.

Bredemeier, C. and Mundstock, C. M. (2000) Regulação da absorção e assimilação do nitrogênio nas plantas, Ciência Rural, 30(2), pp. 365-372.

Bruto, M., Prigent-Combaret, C., Muller, D. and Moënne-Loccoz, Y. (2014) Analysis of genes contributing to plant-beneficial functions in plant growth-promoting rhizobacteria and related Proteobacteria, Sci Rep - UK, 4, pp. 6261.

Burg, S. E. and Thimann, K. V. (1959) The physiology of ethylene formation in apples, Proc. Natl. Acad. Sci., 45, pp. 335-344.

Camilios-Neto, D., Bonato, P., Wassem, R., Tadra-Sfeir, M. Z., Brusamarello-Santos, L. C. C., Valdameri, G., Donatti, L., Faoro, H., Weiss, V. A., Chubatsu, L. S., Pedrosa, F. O. and Souza, E. M. (2014) Dual RNA-seq transcriptional analysis of wheat roots colonized by *Azospirillum brasilense* reveals up-regulation of nutrient acquisition and cell cycle genes, BMC Genomics, 15.

Carraro, N., Forestan, C., Canova, S., Traas, J. and Varotto, S. (2006) ZmPIN1a and ZmPIN1b encode two novel putative candidates for polar auxin transport and plant architecture determination of maize, Plant Physiol, 142(1), pp. 254-64.

Cassán, F., Bottini, R., Schneider, G. and Piccoli, P. (2001) *Azospirillum brasilense* and *Azospirillum lipoferum* Hydrolyze Conjugates of GA20 and Metabolize the Resultant Aglycones to GA1 in Seedlings of Rice Dwarf Mutants1, Plant Physiol, 125, pp. 2053–2058.

Cassán, F., Perrig, D., Sgroy, V., Masciarelli, O., Penna, C. and Luna, V. (2009) *Azospirillum brasilense* Az39 and *Bradyrhizobium japonicum* E109, inoculated singly or in combination, promote seed germination and early seedling growth in corn (Zea mays L.) and soybean (Glycine max L.), Eur J Soil Biol, 45(1), pp. 28-35.

Chen, J. F. and Gallie, D. R. (2010) Analysis of the functional conservation of ethylene receptors between maize and *Arabidopsis*, Plant Mol Biol, 74(4-5), pp. 405-21.

Choi, Y.-J., Aliota, M. T., Mayhew, G. F., Erickson, S. M. and Christensen, B. M. (2014) Dual RNA-seq of Parasite and Host Reveals Gene Expression Dynamics during Filarial Worm–Mosquito Interactions, PLOS Neglect Trop D, 8(5), pp. e2905.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D. J., Elo, L. L., Zhang, X. and Mortazavi, A. (2016) A survey of best practices for RNA-seq data analysis, Genome Biol, 17(1), pp. 13.

Crépin, A., Barbey, C., Beury-Cirou, A., Hélias, V., Taupin, L., Reverchon, S., Nasser, W., Faure, D., Dufour, A., Orange, N., Feuilloley, M., Heurlier, K., Burini, J.-F. and Latour, X. (2012) Quorum Sensing Signaling Molecules Produced by Reference and Emerging Soft-Rot Bacteria (Dickeya and Pectobacterium spp.), *PLOS ONE*, 7(4), pp. e35176.

Daly, P., van Munster, J. M., Kokolski, M., Sang, F., Blythe, M. J., Malla, S., Velasco de Castro Oliveira, J., Goldman, G. H. and Archer, D. B. (2017) Transcriptomic responses of mixed cultures of ascomycete fungi to lignocellulose using dual RNA-seq reveal inter-species antagonism and limited beneficial effects on CAZyme expression, Fungal Genet Biol, 102, pp. 4-21.

Duarte, J. d. O., Garcia, J. C. and Miranda, R. A. (2015) Economia da produção. Sistemas de Produção: Cultivo do Milho. Brasília,DF: Embrapa Milho e Sorgo. Available at: https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemasdeproducaolf6_1ga1ceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&p_r_p_-76293187_sistemaProducaoId=7905&p_r_p_-996514994_topicoId=8658.

Duca, D., Lorv, J., Patten, C. L., Rose, D. and Glick, B. R. (2014) Indole-3-acetic acid in plant–microbe interactions, *Antonie van Leeuwenhoek*, 106, pp. 85-125.

Elmerich, C., De Zamaroczy, M., Arsène, F., Pereg, L., Paquelin, A. and Kaminski, A. (1997) Regulation of NIF Gene Expression and Nitrogen Metabolism in *Azospirillum*, Soil Biol Biochem, 29(5/6), pp. 847-852.

Faleiro, A. L., Pereira, T. P., Espindula, E., Brod, F. C. A. and Arisi, A. C. M. 2013. Real time PCR detection targeting *nifA* gene of plant growth promoting bacteria *Azospirillum brasilense* strain FP2 in maize roots. Symbiosis,61(3), pp. 125-133.

FAO (2016) *FAOSTAT*: United Nations. Available at: http://faostat3.fao.org/home/index.html#VISUALIZE (Accessed: 14/09/2018 2018).

Fehr, W. R., Caviness, C. E., Burmood, D. T. and Pennington, J. S. (1971) Stage of Development Descriptions for Soybeans, *Glycine max* (L.) Merrill1, Crop Sci, 11, pp. 929-931.

Fornasieri Filho, D. (1992) *A Cultura do Milho.*Jaboticabal: FUNEP.

Förstner, K. U., Vogel, J. and Sharma, C. M. (2014) READemption—a tool for the computational analysis of deep-sequencing–based transcriptome data, Bioinformatics, 30(23), pp. 3421-3423.

Hayden, K. J., Garbelotto, M., Knaus, B. J., Cronn, R. C., Rai, H. and Wright, J. W. (2014) Dual RNA-seq of the plant pathogen Phytophthora ramorum and its tanoak host, Tree Genet Genomes, 10(3), pp. 489-502.

Hegedűs, Z., Zakrzewska, A., Ágoston, V. C., Ordas, A., Rácz, P., Mink, M., Spaink, H. P. and Meijer, A. H. (2009) Deep sequencing of the zebrafish transcriptome response to mycobacterium infection, Mol Immunol, 46(15), pp. 2918-2930.

Hirano, K., Ueguchi-Tanaka, M. and Matsuoka, M. (2008) GID1-mediated gibberellin signaling in plants, Trends Plant Sci, 13(4), pp. 192-9.

Hoagland, D. R. and Arnon, D. (1938) The water culture method for growing plants without soil, Circular Califórnia Agricultural Experiment Station, 347(Revised 1950), pp. 32

Hungria, M., Campo, R. J., Souza, E. M. and Pedrosa, F. O. (2010) Inoculation with selected strains of *Azospirillum brasilense* and *A. lipoferum* improves yields of maize and wheat in Brazil, Plant Soil, 331(1-2), pp. 413-425.

Husnik, F. and McCutcheon, J. P. (2017) Functional horizontal gene transfer from bacteria to eukaryotes, Nature Rev Microbiol, 16, pp. 67.

Jahn, C. E., Charkowski, A. O. and Willis, D. K. (2008) Evaluation of isolation methods and RNA integrity for bacterial RNA quantitation, J Microbiol Meth, 75(2), pp. 318-324.

Junior, F. B. d. R., Machado, C. T. d. T., Machado, A. T. and Sodek, L. (2008) Inoculação de *Azospirillum amazonense* em Dois Genótipos de Milho sob Diferentes Regimes de Nitrogênio, *R. Bras. Ci.* Solo, 32, pp. 1139-1146.

Kovalchuk, A., Zeng, Z., Ghimire, R. P., Kivimäenpää, M., Raffaello, T., Liu, M., Mukrimin, M., Kasanen, R., Sun, H., Julkunen-Tiitto, R., Holopainen, J. K. and Asiegbu, F. O. (2019) Dual RNA-seq analysis provides new insights into interactions between Norway spruce and necrotrophic pathogen *Heterobasidionan nosum* s.l, BMC Plant Biol, 19(1), pp. 2.

LaMonte, G. M., Orjuela-Sanchez, P., Calla, J., Wang, L. T., Li, S., Swann, J., Cowell, A. N., Zou, B. Y., Abdel-Haleem Mohamed, A. M., Villa Galarce, Z. H., Moreno, M., Tong Rios, C., Vinetz, J. M., Lewis, N. and Winzeler, E. A. (2019) Dual RNA-seq identifies human mucosal immunity protein Mucin-13 as a hallmark of *Plasmodium* exoerythrocytic infection, Nat Comm, 10(1), pp. 488.

Lanubile, A., Ferrarini, A., Maschietto, V., Delledonne, M., Marocco, A. and Bellin, D. (2014) Functional genomic analysis of constitutive and inducible defense responses to *Fusarium verticillioides* infection in maize genotypes with contrasting ear rot resistance, BMC Genomics, 15(1), pp. 710.

Li, Y., Wang, W., Feng, Y., Tu, M., Wittich, P. E., Bate, N. J. and Messing, J. (2019) Transcriptome and metabolome reveal distinct carbon allocation patterns during internode sugar accumulation in different sorghum genotypes,Plant Biotechnol J, 17(2), pp. 472-487.

Ljung, K. and et al. (2005) Sites and regulation of auxin biosynthesis in *Arabidopsis* roots, Plant Cell., 17, pp. 1090-1104.

Ljung, K., Bhalerao, R. P. and SandberG, G. (2001) Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth, Plant., 1(29), pp. 465-474.

Machado, A. T., Sodek, L., Döbereiner, J. and Reis, V. M. (1998) Efeito da adubação nitrogenada e da inoculação com bactérias diazotróficas no comportamento bioquímico da cultivar de milho Nitroflint, Pesq. Agropec. Bras., 33, pp. 961-970.

Malhotra, M. and Srivastava, S. (2006) Targeted engineering of *Azospirillum brasilense* SM with indole acetamide pathway for indoleacetic acid over-expression, Can J Microbiol, 52(11), pp. 1078-1084.

Malhotra, M. and Srivastava, S. (2008) An ipdC gene knock-out of *Azospirillum brasilense* strain SM and its implications on indole-3-acetic acid biosynthesis and plant growth promotion, Antonie van Leeuwenhoek, 93(4), pp. 425-433.

Mashiguchi, K., Tanaka, K., Sakai, T., Sugawara, S., Kawaide, H., Natsume, M., Hanada, A., Yaeno, T., Shirasu, K., Yao, H., McSteen, P., Zhao, Y., Hayashi, K.-i., Kamiya, Y. and Kasahara, H. (2011) The main auxin biosynthesis pathway in *Arabidopsis*, P Natl Acad Sci USA, 108(45), pp. 18512-18517.

Mateus, I. D., Masclaux, F. G., Aletti, C., Rojas, E. C., Savary, R., Dupuis, C. and Sanders, I. R. (2019) Dual RNA-seq reveals large-scale non-conserved genotype × genotype-specific genetic reprograming and molecular crosstalk in the mycorrhizal symbiosis, ISME J, 13(5), pp. 1226-1238.

Mela, F., Fritsche, K., de Boer, W., van Veen, J. A., de Graaff, L. H., van den Berg, M. and Leveau, J. H. J. (2011) Dual transcriptional profiling of a bacterial/fungal confrontation: *Collimonas fungivorans* versus *Aspergillusniger*, ISME J, 5(9), pp. 1494-1504.

Molina-Favero, C., Creus, C. M., Simontacchi, M., Puntarulo, S. and Lamattina, L. (2008) Aerobic Nitric Oxide Production by *Azospirillum brasilense* Sp245 and Its Influence on Root Architecture in Tomato, Mol Plant Microbe In, 21(7), pp. 1001-1009.

Montoya, D. J., Andrade, P., Silva, B. J. A., Teles, R. M. B., Ma, F., Bryson, B., Sadanand, S., Noel, T., Lu, J., Sarno, E., Arnvig, K. B., Young, D., Lahiri, R., Williams, D. L., Fortune, S., Bloom, B. R., Pellegrini, M. and Modlin, R. L. (2019) Dual RNA-Seq of Human Leprosy Lesions Identifies Bacterial Determinants Linked to Host Immune Response, Cell Rep, 26(13), pp. 3574-3585.e3.

Moon, J., Parry, G. and Estelle, M. (2004) The Ubiquitin-Proteasome Pathway and Plant Development, Plant Cell, 16(12), pp. 3181.

Mutha, N. V. R., Mohammed, W. K., Krasnogor, N., Tan, G. Y. A., Wee, W. Y., Li, Y., Choo, S. W. and Jakubovics, N. S. (2019) Transcriptional profiling of coaggregation interactions between *Streptococcus gordonii* and *Veillonella parvula* by Dual RNA-Seq, Sci Rep - UK, 9(1), pp. 7664.

Nishimura, T., Hayashi, K., Suzuki, H., Gyohda, A., Takaoka, C., Sakaguchi1, Y., Matsumoto, S., Kasahara, H., Sakai, T., Kato, J., Kamiya, Y. and Koshiba1, T. (2014) Yucasin is a potent inhibitor of YUCCA, a key enzyme in auxin biosynthesis,Plant J, 77, pp. 352–366.

Ona, O., Van Impe, J., Prinsen, E. and Vanderleyden, J. (2005) Growth and indole-3-acetic acid biosynthesis of *Azospirillum brasilense* Sp245 is environmentally controlled, FEMS Microbiol Lett, 246(1), pp. 125-32.

Oshlack, A., Robinson, M. D. and Young, M. D. (2010) From RNA-seq reads to differential expression results, Genome Biol, 11(12), pp. 220.

Packard, H., Burke, A. K., Jensen, R. V. and Stevens, A. M. (2017) Analysis of the in planta transcriptome expressed by the corn pathogen *Pantoea stewartii* subsp. stewartii via RNA-Seq, PeerJ, 5(e3237).

Pankievicz, V. C. S., Camilios-Neto, D., Bonato, P., Balsanelli, E., Tadra-Sfeir, M. Z., Faoro, H., Chubatsu, L. S., Donatti, L., Wajnberg, G., Passetti, F., Monteiro, R. A., Pedrosa, F. O. and Souza, E. M. (2016) RNA-seq transcriptional profiling of *Herbaspirillum seropedicae* colonizing wheat (*Triticum aestivum*) roots, Plant Mol Biol, 90(6), pp. 589-603.

Perrig, D., Boiero, M. L., Masciarelli, O. A., Penna, C., Ruiz, O. A., Cassan, F. D. and Luna, M. V. (2007) Plant-growth-promoting compounds produced by two agronomically important strains of *Azospirillum brasilense*, and implications for inoculant formulation, Appl Microbiol Biotechnol, 75(5), pp. 1143-50.

Petrova, O. E., Garcia-Alcalde, F., Zampaloni, C. and Sauer, K. (2017) Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes, Sci Rep - UK, 7, pp. 41114.

Poletto, N. (2004) *Adição de carbono e nitrogênio e sua relação com os estoques no solo e com o rendimento do milho em sistemas de manejo.* Master, Universidade do Rio Grande do Sul, Porto Alegre,RS, Brazil.

Pontiroli, A., Rizzi, A., Simonet, P., Daffonchio, D., Vogel, T. M. and Monier, J. M. (2009) Visual evidence of horizontal gene transfer between plants and bacteria in the phytosphere of transplastomic tobacco, Appl Environ Microbiol, 75(10), pp. 3314-22.

Quispe-Huamanquispe, D. G., Gheysen, G. and Kreuze, J. F. (2017) Horizontal Gene Transfer Contributes to Plant Evolution: The Case of *Agrobacterium* T-DNAs, Front Plant Sci, 8, pp. 2015.

Racolta, A., Bryan, A. C. and Tax, F. E. (2014) The receptor-like kinases GSO1 and GSO2 together regulate root growth in *Arabidopsis* through control of cell division and cell fate specification, Dev Dynam, 243(2), pp. 257-278.

Reeder, S. M., Palmer, J. M., Prokkola, J. M., Lilley, T. M., Reeder, D. M. and Field, K. A. (2017) *Pseudogymnoas cusdestructans* transcriptome changes during white-nose syndrome infections, Virulence, 8(8), pp. 1695-1707.

Reynders, L. and Vlassak, K. (1979) Conversion of tryptophan to indoleacetic acid by *Azospirillum brasilense*, Soil Biol Biochem, 11(5), pp. 547-548.

Ribaudo, C. M., Krumpholz, E. M., Cassán, F. D., Bottini, R., Cantore, M. L. and Curá, J. A. (2006) *Azospirillum* sp. Promotes Root Hair Development in Tomato Plants through a Mechanism that Involves Ethylene,J Plant Growth Regul, 25(2), pp. 175-185.

Salazar-Cerezo, S., Martínez-Montiel, N., García-Sánchez, J., Pérez-y-Terrón, R. and Martínez-Contreras, R. D. (2018) Gibberellin biosynthesis and metabolism: A convergent route for plants, fungi and bacteria, Microbiol Res, 208, pp. 85-98.

Santos, M. F., Padua, V. L. M., Nogueira, E. M., Hemerly, A. S. and Domont, G. B. (2010) Proteome of *Gluconacetobacter diazotrophicus* co-cultivated with sugarcane plantlets, J Proteomics, 73(5), pp. 917-31.

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., Blaxter, M. and Barton, G. J. (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?, RNA, 22(6), pp. 839-51.

Simon, S. and Petrasek, J. (2011) Why plants need more than one type of auxin, Plant Sci, 180(3), pp. 454-60.

Somasegaran, P. and Hoben, H. J. (1994) *Handbook for Rhizobia: Methods in Legume-Rhizobium Technology*. New York, NY: Springer.

Spaepen, S. and Vanderleyden, J. (2010) Auxin and Plant-Microbe Interactions, Cold Spring Harb Perspect Biol, (3 : a001438).

Spaepen, S., Vanderleyden, J. and Remans, R. (2007) Indole-3-acetic acid in microbial and microorganism-plant signaling, FEMS Microbiol Rev, 31(4), pp. 425-448.

Steenhoudt, O. and Vanderleyden, J. (2000) *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects, FEMS Microbiol Rev, 24, pp. 487-506.

Tanimoto, M., Roberts, K. and Dolan, L. (1995) Ethylene is a positive regulator of root hair development in *Arabidopsis thaliana*, Plant Journal, 08, pp. 943-948.

Thakare, D., Yang, R., Steffen, J. G., Zhan, J., Wang, D., Clark, R. M., Wang, X. and Yadegari, R. (2014) RNA-Seq analysis of laser-capture microdissected cells of the developing central starchy endosperm of maize, Genomics Data, 2, pp. 242-245.

Tromas, A. and Perrot-Rechenmann, C. (2010) Recent progress in auxin biology, Comptes Rendus Biologies, 333(4), pp. 297-306.

Tsuwamoto, R., Fukuoka, H. and Takahata, Y. (2008) GASSHO1 and GASSHO2 encoding a putative leucine-rich repeat transmembrane-type receptor kinase are essential for the normal development of the epidermal surface in *Arabidopsis* embryos, Plant J, 54(1), pp. 30-42.

Verna, J.P., Yadav, J., Tiwari, K. N., Singh, L. and Singh, V. (2010) Impact of Plant Growth Promoting Rhizobacteria on Crop Production, Int J Agric Res, 5(11), pp. 954-983.

Verwaaijen, B., Wibberg, D., Kröber, M., Winkler, A., Zrenner, R., Bednarz, H., Niehaus, K., Grosch, R., Pühler, A. and Schlüter, A. (2017) The *Rhizoctonia solani* AG1-IB (isolate 7/3/14) transcriptome during interaction with the host plant lettuce (Lactuca sativa L.), PLOS One, 12(5).

Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics, Nat Rev Genet, 10(1), pp. 57-63.

Weber, M. A. and Mielniczuk, J. (2009) Estoque e Disponibilidade de Nitrogênio no Solo em Experimento de Longa Duração, R. Bras. Ci. Solo, 33, pp. 429-437.

Westermann, A. J., Förstner, K. U., Amman, F., Barquist, L., Chao, Y., Schulte, L. N., Müller, L., Reinhardt, R., Stadler, P. F. and Vogel, J. (2016) Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions, Nature, 529, pp. 496-501.

Westermann, A. J., Gorski, S. A. and Vogel, J. (2012) Dual RNA-seq of pathogen and host, Nat Rev Microbiol, 10, pp. 618-630.

Westermann, A. J. and Vogel, J. (2018) Host-Pathogen Transcriptomics by Dual RNA-Seq, *Bacterial Regulatory RNA: Methods and Protocols, Methods in Molecular Biology*. New York, NY: Humana Press.

Wolf, T., Philipp, K., Brunke, S. and Linde, J. (2018) Twos company: studying interspecies relationships with dual RNA-seq, Curr Opin Microbiol, 42, pp. 7–12.

Won, C., Shen, X., Mashiguchi, K., Zheng, Z., Dai, X., Cheng, Y., Kasahara, H., Kamiya, Y., Chory, J. and Zhao, Y. (2011) Conversion of tryptophan to indole-3-acetic acid by Tryptophan Aminotransferases of *Arabidopsis* and YUCCAs in *Arabidopsis*, P Natl Acad Sci USA, 108(45), pp. 18518.

Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y. (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters, OMICS, 16(5), pp. 284-287.

Yue, J., Hu, X. and Huang, J. (2014) Origin of plant auxin biosynthesis, Trends Plant Sci, 19(12), pp. 764-770.

Zhao, Y. (2010) Auxin biosynthesis and its role in plant development, Annu Rev Plant Biol, 61, pp. 49-64.

Zhao, Y. (2012) Auxin Biosynthesis: A Simple Two-Step Pathway Converts Tryptophan to Indole-3-Acetic Acid in Plants, Mol Plant, 5(2), pp. 334-338.

Zhao, Y. (2014) Auxin Biosynthesis, *The Arabidopsis Book*, pp. e0173.

Zimmer, W., Wesche, M. and Timmermans, L. (1998) Identification and Isolation of the Indole-3-Pyruvate Decarboxylase Gene from *Azospirillum brasilense* Sp7: Sequencing and Functional Analysis of the Gene Locus, Curr Microbiol, 36(6), pp. 327-331.