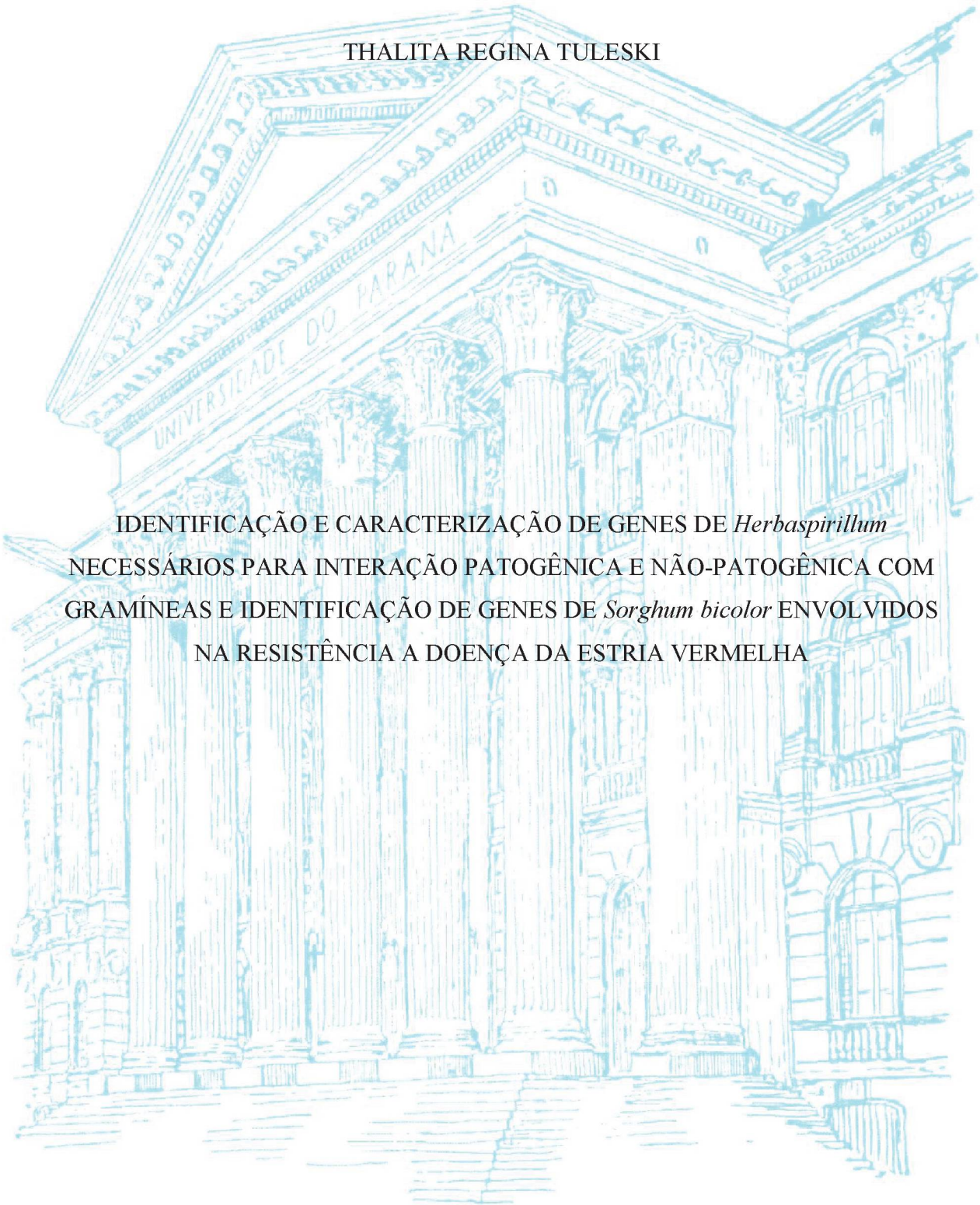
UNIVERSIDADE FEDERAL DO PARANÁ

THALITA REGINA TULESKI



IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE *Herbaspirillum* NECESSÁRIOS PARA INTERAÇÃO PATOGÊNICA E NÃO-PATOGÊNICA COM GRAMÍNEAS E IDENTIFICAÇÃO DE GENES DE *Sorghum bicolor* ENVOLVIDOS NA RESISTÊNCIA A DOENÇA DA ESTRIA VERMELHA

CURITIBA

2017

THALITA REGINA TULESKI

IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE *Herbaspirillum* NECESSÁRIOS PARA INTERAÇÃO PATOGÊNICA E NÃO-PATOGÊNICA COM GRAMÍNEAS E IDENTIFICAÇÃO DE GENES DE *Sorghum bicolor* ENVOLVIDOS NA RESISTÊNCIA A DOENÇA DA ESTRIA VERMELHA

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciências – Bioquímica, no Curso de Pós-Graduação em Bioquímica e Biologia Molecular, Setor de Ciências Biológicas, da Universidade Federal do Paraná.
Orientadora: Prof[a]. Dr[a]. Rose Adele Monteiro
Co-Orientador: Prof. Dr. Emanuel Maltempi de Souza

CURITIBA

2017

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Rosilei Vilas Boas – CRB/9-939).

Tuleski, Thalita Regina
Identificação e caracterização de genes de *Herbaspirillum* necessários para interação patogênica e não-patogênica com gramíneas e identificação de genes de Sorghum bicolor envolvidos na resistência a doença da estria vermelha. / Thalita Regina Tuleski. – Curitiba, 2017.
160 f. : il. ; 30cm.

Orientadora: Rose Adele Monteiro
Coorientador: Emanuel Maltempi de Souza
Tese (Doutorado) – Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Ciências – Bioquímica.

1. *Herbaspirillum*. 2. Bactérias fitopatogênicas. 3. Setaria. 4. Plantas. 5. Genes. I. Título. II. Monteiro, Rose Adele. III. Souza, Emanuel Maltempi de. IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Ciências – Bioquímica.

CDD (20. ed.) 584.92

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS (BIOQUÍMICA) da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **THALITA REGINA TULESKI** intitulada **"IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES DE *Herbaspirillum* NECESSÁRIOS PARA INTERAÇÃO PATOGÊNICA E NÃO-PATOGÊNICA COM GRAMÍNEAS E IDENTIFICAÇÃO DE GENES DE *Sorghum bicolor* ENVOLVIDOS NA RESISTÊNCIA A DOENÇA DA ESTRIA VERMELHA"**. Após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer favorável pela sua APROVAÇÃO.

CURITIBA, 31 de Janeiro de 2017.

ROSE ADELE MONTEIRO
Presidente da Banca Examinadora (UFPR)

MARCO AURELIO S. DE OLIVEIRA
Avaliador Externo (UEM)

DOUMIT CAMILIOS NETO
Avaliador Externo (UEL)

MICHELLE ZIBETTI TADRA SFEIR
Avaliador Interno (UFPR)

FABIO LOPES OLIVARES
Avaliador Externo (UENF)

AGRADECIMENTOS

Agradeço a minha família.

Aos meus pais Sonia e Mauro, pelo suporte incondicional dado a mim desde o começo de tudo.

Ao Gabriel, companheiro de todos os momentos. Obrigada pelo apoio nestes quatro anos de doutorado e por escolher ser minha família e parte da minha vida!

Aos meus irmãos, Aline e Alvaro, ao meu sobrinho Pedro, aos meus avós, tios e primos por todos os momentos de carinho e alegria.

Aos meus orientadores: Professora Rose Adele Monteiro pelo suporte acadêmico, incentivo e carinho do dia-a-dia e Professor Emanuel Maltempi de Souza, pelo constante incentivo e dedicação em formar pessoas cada vez melhores. Ao meu orientador de doutorado sanduíche Dr. Gary Stacey por colaborar com o meu desenvolvimento pessoal e profissional e por ter me dado a oportunidade de conviver com pessoas únicas durante meu tempo no Stacey's Lab.

A banca interna, Dra Michelle Zibetti Tadra-Sfeir e Professor Leonardo Magalhães Cruz por acompanhar este projeto durante os 4 anos e contribuir para o desenvolvimento de cada passo.

Aos técnicos do laboratório, Valter, Roseli Prado, Lucinéia e Marilza pelo suporte para a realização de cada experimento.

A todos os professores do departamento de Bioquímica e Biologia Molecular, em especial aos professores do Grupo de Fixação de Nitrogênio, por contribuir ativamente para a formação de cada aluno.

Aos amigos e a todos os colegas de laboratório, pela companhia e suporte no dia-a-dia.

Obrigada

RESUMO

O gênero *Herbaspirillum* compreende algumas espécies bactérias capazes de fixar o nitrogênio atmosférico em condições de microaerofilia. Várias espécies de *Herbaspirillum* já foram descritas, entre elas estão *Herbaspirillum seropedicae* e *Herbaspirillum rubrisubalbicans*. *Herbaspirillum rubrisubalbicans* é um dizotrofo, encontrado em associação benéfica com várias plantas de interesse econômico, entretanto quando se associa a variedade B4362 de cana de açúcar ou a variedades susceptíveis de sorgo é responsável por causar a doença da estria mosqueada e estria vermelha, respectivamente. Estas doenças diminuem a área fotossintética e a vida útil da folha devido aos danos causados nos tecidos foliares: manchas e necrose. Alguns fatores, como o Sistema de Secreção do Tipo III já foram descritos como essenciais para o desenvolvimento da doença, porém, outros fatores ainda precisam ser elucidados para o melhor entendimento dos processos de interação planta-bactéria envolvendo este microrganismo. Neste trabalho, a biossíntese de celulose, um exopolissacarídeo (EPS) sintetizado por *H. rubrisubalbicans* e ausente em *H. seropedicae*, foi avaliada quanto ao seu envolvimento com os processos de colonização, formação de biofilme e patogenicidade. O genótipo da planta parece ter papel crucial para o desenvolvimento da doença, técnicas de RNA-seq e QTL (Quantitative trait locus) foram aplicadas em genótipos de sorgo com o objetivo de identificar fatores, presentes na planta, potencialmente envolvidos com a interação entre *H. rubrisubalbicans* e Sorgo. Dentre os resultados obtidos, genes envolvidos com as vias clássicas de interação planta-patógeno foram identificados e sugerem *H. rubrisubalbicans* como um bom modelo para o estudo da patogenicidade em plantas de sorgo. Já *H. seropedicae* promove uma associação exclusivamente benéfica durante a interação com gramíneas, sendo responsável por aumentar até 120% do nitrogênio acumulado. Dentre as várias gramíneas que se associam a *H. seropedicae* está *Setaria viridis.* A associação entre *Setaria viridis* e *H. seropedicae* (em conjunto com *Azospirillum brasilense*) promove aumento significativo no tamanho da raiz e no número de raízes laterais. Neste trabalho também foram analisados os fatores relacionados a interação benéfica entre *Setaria viridis* e *H. seropedicae* através da técnica de *TnSeq*, bem como o comportamento desta bactéria em diferentes condições de cultivo. Diversos genes candidatos foram identificados e mostram possíveis vias pelas quais planta e bactéria interagem. Sendo estas duas espécies de *Herbaspirillum* geneticamente relacionadas, o entendimento das diferenças durante a interação planta bactéria torna-se importante para que os mecanismos envolvidos sejam elucidados e para que o potencial destes organismos como promotores de crescimento vegetal seja totalmente aproveitado.

Palavras chave: *Herbaspirillum rubrisubalbicans*, *Herbaspirillum seropedicae*, Interação planta-bactéria, *Sorghum bicolor*, patogenicidade, associação benéfica.

ABSTRACT

The *Herbaspirillum* genus comprises some nitrogen fixing bacteria, capable of fixate the atmospheric nitrogen under microaerophilic conditions. Several species of *Herbaspirillum* have been described, among them *Herbaspirillum seropedicae* and *Herbaspirillum rubrisubalbicans*. *Herbaspirillum rubrisubalbicans* is a diazotroph, being found in beneficial association with many plants of economic interest, however when the association occurs with sugarcane B4362 or a susceptible sorghum genotypes, the bacteria is responsible for causes the mottled stripe disease and red stripe disease respectively. These diseases decrease the photosynthetic area and a useful life of the leaf due to the damages caused on the foliar tissues: stripes and necrosis. Some factors, such as the Type III Secretion System have been already described as essential to the development of the disease; however, other factors still need to be elucidated for a better understanding of the plant-bacterial interaction processes. In this work, the cellulose biosynthesis, an exopolysaccharide (EPS) synthesized by *H. rubrisubalbicans* and absent in *H. seropedicae*, was evaluated for its involvement with colonization, biofilm formation and pathogenicity. As the plant genotype is also crucial to the development of the disease, the techniques of RNA-seq and QTL (Quantitative trait locus) were applied in sorghum genotypes to identify the factors present in the plant, potentially involved in the interaction between *H. rubrisubalbicans* and Sorghum. Among the results obtained, genes involved with classical pathways of plant-pathogen interaction were identified and suggested *H. rubrisubalbicans* as a good model for the study of pathogenicity in Sorghum plants. *H. seropedicae* promotes an exclusively beneficial association during the interaction with grasses, being responsible for increasing up to 120% of the accumulated nitrogen. Among the several grasses associated with *H. seropedicae* are *Setaria viridis*. The association between *Setaria viridis* and *H. seropedicae* (together with *Azospirillum brasilense*) promotes a significant increase in root size and number of lateral roots. In this work we also analyzed the factors related to the beneficial interaction between *Setaria viridis and H. seropedicae* through the *Transposon Sequencing* technique (TnSeq), as well as the behavior of this bacterium under different culture conditions. Several possible genes were identified and showed possible interaction pathways between plant and the bacteria. Being these two species genetically related, the understanding of the differences during the interaction becomes important to elucidate the mechanisms involved and to the potential of this organisms as plant growth promoters is fully utilized.

Key-words: *Herbaspirillum rubrisubalbicans*, *Herbaspirillum seropedicae,* Plant-bacteria Interaction, *Sorghum bicolor*, pathogenicity, beneficial association.

LISTA DE FIGURAS

LISTA DE FIGURAS SUPLEMENTATRES

LISTA DE TABELAS

LISTA DE TABELAS SUPLEMENTARES

SUMARIO

CAPITULO I - INTRODUÇÃO E REVISÃO BIBLIOGRÁFICA

1.1 ASSOCIAÇÃO ENTRE GRAMÍNEAS E ORGANISMOS DIAZOTRÓFICOS ENDOFÍTICOS DO GÊNERO *Herbaspirillum* sp.

Alguns microrganismos, como as rizobactérias, são capazes de promover o crescimento das plantas através de diferentes mecanismos, tais como fixação de nitrogênio e produção de fitohormonios (RHICHARDSON et al., 2009). O gênero *Herbaspirillum* sp. pertence ao grupo das β-proteobactérias e compreende microrganismos gram-negativos, geralmente vibrióides, com 1 a 3 flagelos e com diferentes estilos de vida. Algumas espécies vivem em ambientes aquáticos, no solo ou podem ser encontradas como patógenos oportunistas em doenças humanas (BERG et al., 2005). Em contraste, algumas espécies podem fixar o nitrogênio atmosférico, em condições microaeróbicas, como *Herbaspirillum seropedicae* e *Herbaspirillum rubrisubalbicans*, e associar-se beneficamente com diferentes culturas vegetais, fixando o nitrogênio atmosférico e tornando-o assimilável para planta. Entre as espécies descritas, *H. seropedicae, H. frisingense e H. rubrisubalbicans* são os únicos capazes de fixar nitrogênio atmosférico. *H. seropedicae* e *H rubrisubalbicans* foram encontrados em associação benéfica com gramíneas, aumentando o crescimento e produtividade de várias culturas, como o milho (*Zea mays*), cana-de-açúcar (*Saccharum officinarum*) e arroz (*Oryza sativa*) (BALDANI et al., 1986; BALDANI et al., 1992; BODDEY et al., 1995; JAMES et al., 2000; JAMES et al., 2002; ELBELTAGY et al., 2001 e RONCATO-MACCARI et al., 2003).

A interação entre plantas e bactérias diazotróficas tem sido amplamente estudada, bem como o uso de bactérias endofíticas, como *Azospirillum, Gluconacetobacter, Azoarcus e Herbaspirillum* como promotoras do crescimento vegetal (MONTEIRO et al., 2012). Esse tipo de associação tem sido descrito como importante já que estes organismos podem colonizar diferentes tecidos, promover o crescimento das plantas pelo controle de fitopatógenos, aumento da captação de minerais e a produção de diferentes fitohormônios, além da fixação de nitrogênio (ELBELTAGY et al., 2001; STURZ et al., 2000).

1.1.1 Herbaspirillum seropedicae

Entre os organismos fixadores de nitrogênio conhecidos encontra-se *Herbaspirillum seropedicae* (BALDANI et al., 1986, OLIVARES et al., 1997). *Herbaspirillum seropedicae* é encontrado em associação com várias culturas economicamente importantes, como milho (*Zea mays*), arroz (*Oryza sativa*), sorgo (*Sorghum bicolor*) e cana-de-açúcar (*Saccharum officiarum*) (BALDANI et al., 1986; BALDANI et al., 1992; BODDEY et al., 1995; JAMES et al., 2000; JAMES et al., 2002; ELBELTAGY et al., 2001; GYANESHWAR et al., 2002 e RONCATO-MACCARI et al., 2003). A associação com *H. seropedicae* é normalmente benéfica para a planta, aumentando o Nitrogênio acumulado em até 120% e, consequentemente, aumentando o crescimento da planta (BALDANI et al, 2000).

A associação entre *H. seropedicae* e plantas ainda é mal compreendida, mas alguns fatores importantes já foram elucidados, como a importância da biossíntese de polissacarídeos, a alteração no envelope celular bacteriano (síntese de LPS e peptidoglicanos) e o metabolismo do PHB durante a interação com milho e trigo (BALSANELLI et al., 2015 e PANKIEVICZ et al., 2016).

A interação entre *H.seropedicae* (co-inoculado com *Azospirillum brasilense*) e *Setaria viridis* também mostrou uma elevada colonização bacteriana, bem como um aumento significativo no número de sementes, peso seco/fresco, comprimento da raiz e número de raízes laterais (40 dias após a inoculação) (PANKIEVICZ et al., 2015).

O mecanismo de interação *Herbaspirillum*-planta se inicia pela adesão da bactéria à superfície radicular, colonização dos pontos radiculares secundários e descontinuidades da epiderme seguida pela penetração e dispersão das bactérias nos espaços intercelulares de raízes, aerênquima, tecidos vasculares e aéreos (FIGURA 1-1) (JAMES et al., 1997 e RONCATO-MACCARI et al., 2003).

Os padrões de interação planta bactéria entre *H. seropedicae* e gramíneas indica o grande potencial desta bactéria como um biofertilizante nitrogenado, e a compreensão total da biologia desta bactéria é importante como a elucidação de todas as vias e exigências para o crescimento deste organismo.

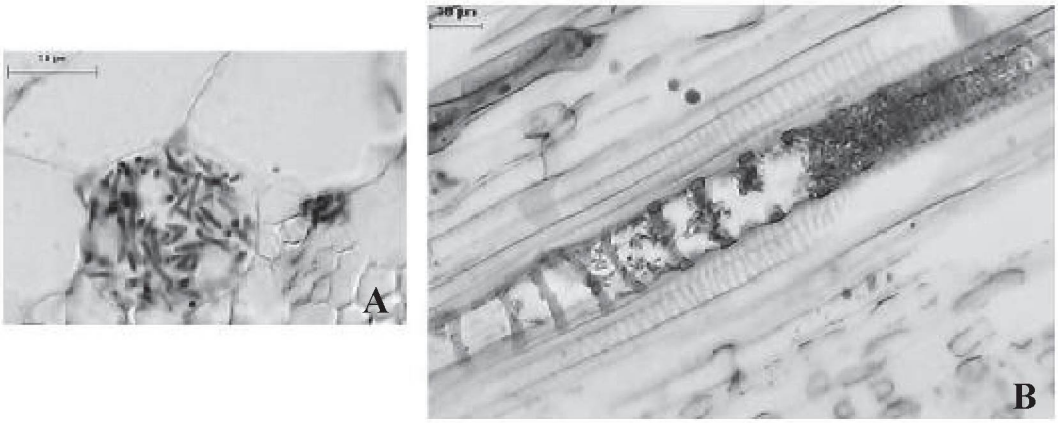


FIGURE 1-1 – Colonização de tecidos de milho e trigo por *H. seropedicae.* (A) *H. seropedicae* colonizando os vasos do xilema de milho 15 dias após a inoculação e (B) Corte longitudinal dos vasos do xilema de trigo mostrando a colonização por *H. seropedicae*, após 15 dias de inoculação. Fonte: RONCATO-MACCARI et al., 2003.

1.1.2 Herbaspirillum rubrisubalbicans

Herbaspirillum rubrisubalbicans é também um organismo diazotrófico endofítico e pode ser isolado de folhas e raízes de gramíneas saudáveis colonizando os tecidos internos e promovendo o crescimento vegetal (BALDANI et al., 1992; OLIVARES et al., 1996 e CRUZ et al., 2001). Assim como em *H. seropedicae*, acredita-se que a interação com plantas inicia-se com a adesão das bactérias à superfície da raiz, seguida da colonização dos pontos de emergência das raízes secundárias ou fissuras na epiderme, seguida da invasão e estabelecimento endofítico nos espaços intercelulares, xilemas, e células de parênquima (JAMES et al., 1997, OLIVARES et al., 1997 e MONTEIRO et al., 2012).

Em contraste com *H. seropedicae, H. rubrisubalbicans* M1 apresenta um comportamento fitopatogênico quando associado ao cultivar B-4362 de cana-de-açúcar e a algumas variedades susceptíveis de sorgo, desenvolvendo, respectivamente, a doença da estria mosqueada e da estria vermelha (FIGURA 1-2) (PIMENTEL et al., 1991 e HALE e WIKIE, 1972). Os sintomas destas doenças são caracterizados por manchas vermelhas e por pontos de necrose tecidual próximos aos pontos de inoculação e ao longo das folhas. Estes sintomas são provavelmente devidos à alta colonização dos tecidos internos, conforme descrito por Olivares (1997). A estria mosqueada aumenta desde o ponto de inoculação em direção ao topo da folha, à medida que a bactéria avança no interior dos vasos da folha (OLIVARES et al., 1997). A estirpe *H rubrisubalbicans* M1 coloniza intensamente o meta e o

protoxilema, bloqueando estes vasos, o que leva à necrose e, consequentemente, diminuindo os níveis fotossintéticos de plantas infectadas (OLIVARES et al., 1997; JAMES et al., 1997). As bactérias encontradas no xilema e no mesófilo são cercadas por uma matriz mucilaginosa, sugerindo a presença de exopolissacarídeos (EPS) e lipopolissacarídeos (LPS) envolvidos com o estabelecimento endofítico deste organismo no hospedeiro (OLIVARES et al., 1997).

FIGURE 1-2- Doença da estria mosqueada causada por *Herbaspirillum rubrisubalbicans* M1 em cana-de-açúcar B-4362. A figura mostra sintomas típicos da doença da estria mosqueada em cana de açúcar. A estrela indica o ponto de inoculação da bactéria, nota-se o desenvolvimento de estrias vermelhas e necrose do tecido. As flechas indicam os sintomas da doença se manifestando em regiões acima do ponto de inoculação, indicando o avanço das bactérias no interior dos vasos da folha. (FONTE: OLIVARES et al., 1997)

Herbaspirillum rubrisubalbicans pode ser encontrado colonizando os tecidos internos do hospedeiro, tais como cavidades substomáticas, espaços intercelulares e xilema (OLIVARES et al., 1995), as bactérias também podem ser encontradas em diferentes áreas da folha (longe do ponto de inoculação) sugerindo uma translocação livre das bactérias no hospedeiro, provavelmente através dos feixes vasculares (PIMENTEL et al., 1991; BALDANI et al., 1992; DOBEREINER et al., 1993 e 1994).

Um grande número de células de *H. rubrisubalbicans* pode ser encontrado no interior dos tecidos das folhas, dentro dos vasos do protoxilema, associado a lacunas, e aparentemente bloqueando os vasos completamente. Já a interação com *H. seropedicae* apresenta um número controlado de bactérias (FIGURE 1-3) (JAMES et al., 1997).

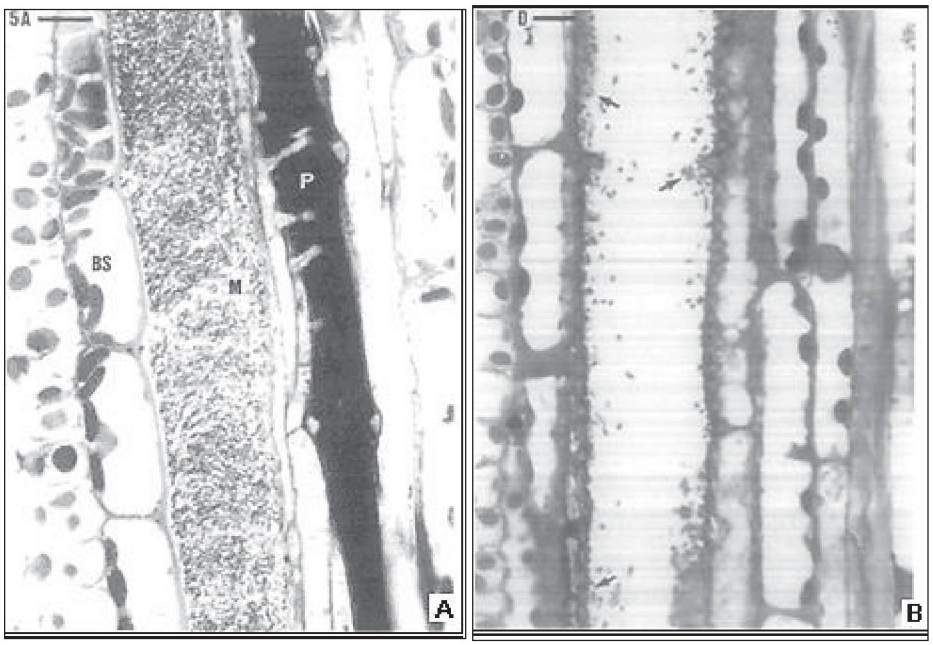


FIGURE 1-3 – *Herbaspirillum rubrisubalbicans* (A) e *Herbaspirillum seropedicae* (B) colonizando feixes vasculares de folha de sorgo. Seções retiradas de áreas de folhas de sorgo mostrando sintomas da doença após 14 dias de inoculação com *Herbaspirillum rubrisubalbicans*, A- seção longitudinal mostrando o xilema e o floema (P) fortemente colonizados por bactérias, onde M = metaxilema e BS = bainha de células. Em B, seção longitudinal de protoxilema de uma folha de sorgo 14 dias após a inoculação com *Herbaspirillum seropedicae*. As bactérias estão concentradas nas paredes dos vasos (setas) e o lúmen do vaso permanece sem células bacterianas (JAMES et al., 1997).

As doenças causadas por *H. rubrisubalbicans* podem reduzir em até 50% a vida útil das folhas, porém, mesmo durante a associação patogênica a bactéria é capaz de expressar os genes da nitrogenase e fixar o nitrogênio atmosférico (JAMES et al., 1997; OLIVARES et al., 1997).

Em 1980, Galli e colaboradores encontraram a variedade de cana de açúcar suscetível B-4362 em Barbados e em 2010 a doença da estria mosqueada foi também descrita em cana-de-açúcar na China (TAN et al., 2010). No entanto, de acordo com Pimentel (1991) e Olivares (1997) não há relato de propagação da doença da estria vermelha/mosqueada nos campos de cana-de-açúcar e sorgo no Brasil, sendo os genótipos agronomicamente importantes resistentes a esta doença. Olivares (1997) também mostrou que a inoculação artificial em um cultivar de cana-de-açúcar amplamente cultivado no Brasil (SP 70-1143) não mostrou qualquer sintoma da doença após a inoculação.

A interação entre *Herbaspirilum rubrisubalbicans* e sorgo é desconhecida, especialmente aos fatores relacionados à planta. Sendo o *Sorghum bicolor* uma importante

gramínea, com potencial para produção de bioenergia, é importante compreender o mecanismo envolvido nesta interação. Bem como é importante estabelecer um modelo para o estudo do sistema imune de sorgo, uma vez que esta planta tem um grande potencial para enfrentar a crise energética e sua compreensão é um passo crucial para a sua utilização como um recurso energético.

1.2 Sorghum bicolor

O sorgo (*Sorghum bicolor*) é a quinta variedade vegetal mais plantada atualmente e é uma importante gramínea com recursos que a tornam um grande candidato para a produção de bioenergia (ROONEY et al., 2007). Entre as características importantes do sorgo estão o potencial de produção agrícola, o sistema de produção já estabelecido, as características de crescimento rápido, o alto teor de carbono e o grande potencial para melhoramento genético. Outra característica crucial é a produção de sorgo poder ser associada a ambientes quentes e secos, tendo uma alta eficiência na utilização da água e sendo altamente tolerância à seca (BOYER et al., 1982; WIEDENFELD et al., 1984 e ROONEY et al., 2007). Combinada com todas essas características, a crise energética aumenta o interesse em explorar o desenvolvimento do sorgo e disponibilizar seus recursos para a produção de bioenergia (McBEE et al., 1987).

1.3 FATORES ENVOLVIDOS NA INTERAÇÃO PLANTA BACTÉRIA ENTRE *HERBASPIRILLUM SP.* E GRAMÍNEAS

As células bacterianas quando aderidas às raízes das plantas produzem exopolissacarídeos e lipopolissacarídeos para ancorar as bactérias na planta, produzindo uma matriz estável e um biofilme que estabiliza a interação planta-bactéria. Os exopolissacáridos (EPS) são polissacarídeos extracelulares produzidos por bactérias (KANG et al., 1979), que podem estar relacionados, entre outras coisas, com os sistemas de secreção, regulação de genes, interação célula-célula, patogênese e simbiose. A celulose é um dos EPS envolvidos nos processos de interação planta bactéria e a importância da sua biossíntese foi descrita em algumas bactérias associativas de plantas, como *Rhizobium leguminosarum* e *Agrobacterium tumefaciens*, demonstrando que a celulose é importante para ancoramento e promoção de uma ligação eficiente e forte das bactérias às raízes (MATTHYSSE et al., 1981; AUSMEES et al., 1999). A biossíntese de celulose em um mutante no operon de biossíntese de celulose *bcs*

(genes *bcsC* e *bcsE*) em *Salmonella enteritidis*, produz um biofilme frágil, instável e facilmente removido (SOLANO et al., 2002).

O biofilme é um dos principais fatores envolvidos nos mecanismos de interação planta-bactéria. A estrutura tridimensional do biofilme compreende agregados bacterianos, imersos em uma matriz extracelular composta principalmente de água, polissacarídeos e proteínas (JAHN et al., 2000). Esta estrutura é eficaz na promoção da sobrevivência bacteriana em diferentes ambientes, aumentando a resistência aos antibióticos, desenvolvendo um metabolismo integrado e facilitando a transferência horizontal de genes (WATNICK e KOLTER et al., 2000, FUQUA et al., 2007).

A formação de biofilme foi descrita como conferindo uma importante vantagem aos organismos patogênicos, sendo considerada o primeiro passo para o estabelecimento eficiente da patogênese. *Pseudomonas aeruginosa* produz biofilme em torno de raízes de *Arabidopsis thaliana*, matando a planta em poucos dias (WALKER et al., 2004). *Azospirillum brasilense*, uma bactéria diazotrófica e promotora do crescimento de plantas forma micro-colônias ou agregados na zona de alongamento e nos pêlos radiculares das raízes do trigo, apontando a importância deste mecanismo também para a interação benéfica com plantas (ASSMUS et al., 1995).

Uma comparação genômica mostra que *H. rubrisubalbicans* M1 tem um *operon* de genes de biossíntese de celulose (*bcs*), que está ausente do genoma de *H. seropedicae* SmR1 (MONTEIRO et al., 2012). Um mutante neste *operon*, no gene *bcsZ*, foi construído por Monteiro (2012) e este mutante demostrou possuir uma padrão de colonização menor em milho (*Zea mays*) quando comparado ao selvagem M1. Uma vez que se demonstrou que a celulose é importante para a interação planta bactéria, torna-se interessante elucidar seu papel também na formação de biofilme por *H. rubrisubalbicans*.

Diversos outros mecanismos ainda devem ser elucidados para a melhor compreensão dos processos de Interação planta bactéria envolvendo a associação benéfica promovida por *H. seropedicae* e *H. rubrisubalbicans* em gramíneas, e a associação patogênica específica promovida por *H. rubrisubalbicans* em sorgo e cana-de-açúcar.

Para uma melhor compreensão do processo de interação das bactérias, uma das técnicas mais poderosas a ser usada, juntamente com o RNASeq, é o TnSeq (Transposon Sequencing). Esta abordagem baseia-se na utilização de transposons, elementos genéticos móveis, para se obter mutagênese aleatória por inserção. Combinando esta técnica com as plataformas de sequenciamento de nova geração é possível identificar e acompanhar as regiões mutagenizadas e seus respectivos genes e funções e assim inferir a importância dos

genes para a sobrevivência da bactéria nas diferentes condições testadas através de ensaios de *fitness* e do cálculo do *gene fitness* para cada gene bacteriano (OPIJNEN e CAMILLI, 2013).

Entre as diferentes metodologias utilizadas para realizar o TnSeq está o Bar-Seq (FIGURA 1-4). Esta metodologia permite associar a mutagênese aleatória (Tn-Seq) ao uso de BarCodes - sequências de DNA - previamente associadas a mutações específicas. Esta correlação funciona como um marcador genético para cada estirpe em um único conjunto, com milhares de mutantes agrupados. Desta forma, é possível estimar a importância de cada gene para condições específicas através da contagem do número de barcodes que correspondem às mutações gênicas mapeadas. Esta técnica é uma ferramenta importante para estudar a relação entre os genes e suas funções em uma variedade de ambientes (WETMORE et al., 2015).

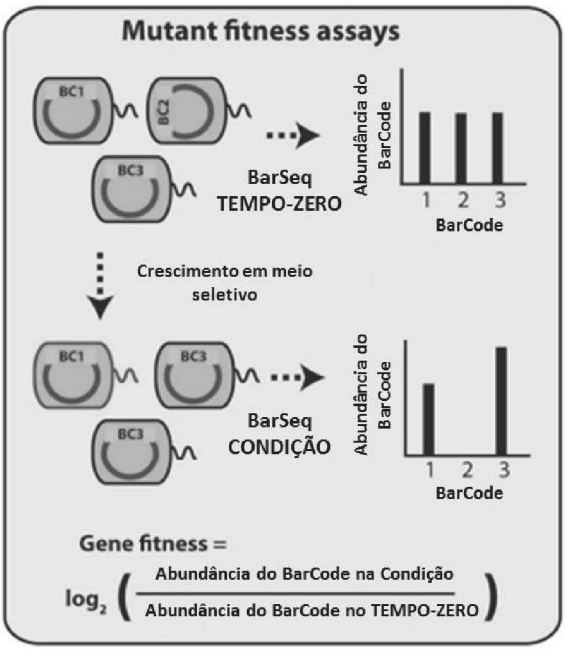


FIGURE 1-4 – Modelo esquemático para realização dos ensaios de Fitness. A técnica de BarSeq se baseia na inserção por transposon de BarCodes de DNA no genoma do organismo de estudo, gerando milhares de mutantes aleatórios identificáveis pelo barcode adicionado. O ensaio de Fitness é baseado na contagem destes barCodes de DNA antes da realização do experimento (BarSeq TEMPO-ZERO) e depois do crescimento na condição a ser analisada (BarSeq CONDIÇÃO). O gene fitness é gerado através da fórmula indicada na Figura e seu valor permite inferir a importância do gene para a sobrevivência do organismo na condição analisada.

Vários trabalhos com microrganismos já foram desenvolvidos utilizando a técnica Tn-Seq, incluindo estudos com organismos patogênicos, como *Vibrio cholerae* e *Pseudomonas aeruginosa*, onde foram encontrados fatores importantes para a patogênese como proteínas efetoras do Sistema de Secreção Tipo VI em *V. cholerae* (DONG et al., 2012 e WETMORE et al., 2015).

Outra importante utilização possível para os ensaios de TnSeq, além do estudo global da exigência bacteriana frente a diferentes ambientes, é anotação de genes com função desconhecida, um dos maiores desafios nos estudos genômicos.

Wetmore (2015) mostra uma alta eficiência e reprodutibilidade dos resultados em um experimento usando mais de cem diferentes fontes de carbono em 5 espécies de bactérias: *Escherichia coli, Phaeobacter inibidores, Pseudomonas stutzeri, Shewanella amazonensis e Shewanella oneidensis*. Este trabalho pôde identificar mais de 5 mil diferentes genes importantes e dependentes da fonte de carbono utilizadas para o crescimento nestas 5 bactérias usando como indicativo o *gene fitness*.

Desta forma, o emprego desta técnica pode auxiliar no entendimento metabólico e associativo de bactéria do gênero *Herbaspirillum* e gramíneas.

CAPITULO II – OBJETIVOS

2.1 OBJETIVO GERAL

Este trabalho tem como objetivo geral identificar possíveis genes relacionados com a interação planta bactéria usando como modelo de estudo bactérias do gênero *Herbaspirillum*: *Herbaspirillum seropedicae* e *Herbaspirillum rubrisubalbicans* e gramíneas, bem como identificar genes importantes de *H. seropedicae* SmR1 relacionados ao crescimento em diferentes condições de cultivo.

2.2 OBJETIVOS ESPECÍFICOS

- Analisar *a função* do operon *bcs* de *H.rubrisubalbicans*, relacionado a biossíntese de celulose, durante a formação de biofilme e Interação planta bactéria;
- Realizar um *screnning* de genótipos de sorgo resistentes e susceptíveis a doença da estria vermelha causada por *H. rubrisubalbicans;*
- Analisar os fatores envolvidos durante a interação patogênica entre *Herbaspirillum rubrisubalbicans* e Sorgo utilizando técnicas de microscopia, QTL e RNASeq;
- Analisar o envolvimento do Sistema Imune Vegetal durante o desenvolvimento da doença da estria vermelha em Sorgo;
- Identificar genes envolvidos durante a associação benéfica entre *Herbaspirillum seropedicae* e *Setária viridis* utilizando a técnica de TnSeq;
- Identificar fatores de *Herbaspirillum seropedicae* SmR1 importantes durante o crescimento em diferentes fontes de Carbono, Nitrogênio e Estresse osmótico.

CAPÍTULO III

Manuscrito ainda não submetido

INVOLVEMENT OF CELLULOSE BIOSYNTHESIS GENES IN BIOFILM FORMATION BY THE BACTERIA *Herbaspirillum rubrisubalbicans* M1

Involvement of cellulose biosynthesis genes in early biofilm formation by the bacteria *Herbaspirillum rubrisubalbicans* M1

ABSTRACT

Three species of the *Herbaspirillum* genus are nitrogen fixers capable of endophytic association with important agricultural crops, such as maize, rice, wheat, sorghum and sugar-cane. *Herbaspirillum rubrisubalbicans,* besides being a diazotroph, causes the mottled stripe disease in susceptible sugar-cane cultivars and also the red stripe disease in some sorghum cultivars. *H. rubrisubalbicans* has a cluster of eight genes involved in cellulose biosynthesis (*bcs*). A mutant strain of *H. rubrisubalbicans* in the gene *bcsZ,* that codes for an 1,4 endoglucanase, showed a decrease in early biofilm formation ability, and in endophytic and epiphytic maize root colonization when compared to wild type M1. The data present in this paper suggest that cellulose biosynthesis is somehow involved in the biosynthesis of biofilm and may affect plant-bacteria interactions.

INTRODUCTION

Plant Growth promoting rhizobacteria (PGPR) can enhance plant growth by different mechanisms such as nitrogen fixation, and phytohormones production (RICHARDSON et al., 2009). The *Herbaspirillum* genus belongs to the Betaroteobacteria class and comprises some species capable of fixing nitrogen and diazotrophic growth. Among the described species, *H. seropedicae, H. frisingense* and *H. rubrisubalbicans* are the only diazotrophs. In this paper we study the capacity of *H. rubrisubalbicans* in synthesize biofilm and its role of the biofilm in plant-bacteria interactions.

Herbaspirillum rubrisubalbicans is capable of colonizing inner tissues and to promote plant growth. The interaction between *H. rubrisubalbicans* and root plants begins with the attachment of the bacteria to the root surface followed by the colonization of emergence points of secondary roots or through crack in the epidermis, followed by invasion and endophytic establishment in intercellular spaces, xylem vessels and parenchyma cells (JAMES et al., 1997, OLIVARES et al., 1997 e MONTEIRO et al., 2012).

H. rubrisubalbicans is phytopathogenic to the sugar cane cultivar B-4362 and to some susceptible sorghum cultivars, developing the mottled stripe and the red stripe diseases, respectively (PIMENTEL et al., 1991 and HALE and WIKIE, 1972). The symptoms,

characterized by red patches in the inoculation points and along the leaves, are probably due to the massive colonization of the internal tissues (OLIVARES et al., 1997). The *H rubrisubalbicans* M1 strain intensively colonizes meta and protoxylem, blocking these vessels, leading to leave necrosis and, consequently, decreasing the photosynthetic levels (OLIVARES et al., 1997; JAMES et al., 1997). The bacteria found in xylem and mesophyll was surrounded by mucus that was suggested to contain exopolysaccharides and lipopolysaccharides (OLIVARES et al., 1997).

Exopolysaccharides (EPS) are extracellular polysaccharides produced by bacteria (Kang et al., 1979). They can be related to secretion systems, gene regulation, cell-cell interactions, pathogenesis and symbiosis. Cellulose is an important EPS in the plant-bacteria interaction, already described in associative organisms such as *Rhizobium* sp. (SMIT et al., 1987). Bacterial cells attached to plant roots produce EPS to anchor the bacteria to the plant, producing a stable matrix and biofilm that stabilize the plant-bacteria interaction (PIMENTEL et al., 1991; BALDANI et al., 1992, HALE and WIKIE, 1972 and PIMENTEL et al., 1991).

A genomic comparison shows that *H. rubrisubalbicans* M1 has an operon of cellulose biosynthesis (*bcs/wss*) genes, that is absent in *H. seropedicae* SmR1 genome (MONTEIRO et al., 2012). A mutant in this operon, *bcsZ* gene, was constructed by Monteiro (2012) and this mutant was impaired in attachment when compared to the wild type strain. The cellulose biosynthesis has been described in some plant associative bacteria, like *Rhizobium leguminosarum* and *Agrobacterium tumefaciens*, and cellulose seems to be important to anchor de bacteria to the roots and promote an efficient and strong attachment (MATTHYSSE et al., 1981; AUSMEES et al., 1999). In *Salmonella enteritidis*, a mutant in the *bcs* operon (*bcs*C and *bcs*E genes) produces a fragile, unstable and easily removed biofilm (SOLANO et al., 2002).

Biofilm is one of the factors knowingly involved in the plant-bacteria interaction mechanisms. The tridimensional structure of biofilm comprises bacterial aggregates, immersed in an extracellular matrix composed mainly of water, polysaccharides and proteins (JAHN et al., 2000). This structure is effective in promoting bacterial survival in different environments increasing antibiotic resistance, developing an integrated metabolism and facilitating horizontal gene transfer (WATNICK and KOLTER, 2000, FUQUA et al., 2007).

The formation of biofilm has been described as conferring an important advantage to pathogenic organisms, being considered the first step to the efficient establishment of pathogenesis. *Pseudomonas aeruginosa* produces biofilm around *A. thaliana* roots, killing the

plant in a few days (WALKER et al., 2004). *Azospirillum brasilense*, a diazotrophic and plant growth promoting bacteria form micro-colonies or aggregates in the elongation zone and in root hairs of wheat roots pointing to the importance of this mechanism to plant bacterial interaction (ASSMUS et al., 1995).

Ausmees et al (1999) and Zevenhulzen et al. (1986) showed a correlation between congo red dye and cellulose content in bacterial cultures. Mutants of *Rhizobium leguminosarum* bv. *trifolii* (AUSMESS et al., 1999) and *Gluconoacetobacter hansenii* (MOHITE and PATIL, 2014) impaired in cellulose production also showed a reduction in congo red binding. A mutation in the *K. xylinus bcsZ* gene showed a decrease in cellulose production and an irregular packing of its fibrils was observed (NAKAI et al., 2013).

Since it has been shown that cellulose is important to biofilm formation in many bacterial species, and that a mutation in the *bcs* operon of *H. rubrisubalbicans* promotes a decrease in its attachment to maize roots, this paper analyzes whether a change in the rate of early biofilm formation affects the extent of maize root colonization by this pathogenic diazotroph.

RESULTS AND DISCUSSION

Production of cellulose by *H. rubrisubalbicans* strains

The *bcs* and *cel* genes are involved in cellulose biosynthesis in many bacteria. Mutation in these genes in *Pseudomonas fluorescens*, *Pseudomonas syringae* and *Gluconacetobacter (Acetobacter) xylinus* affect cellulose production (SPIERS et al., 2003; ROMLING et al., 2002; WONG et al., 1990). Cellulose production by *H. rubrisubalbicans* M1 and TRT1, a mutant in *bcsZ* (previously named *wssD*) gene involved in cellulose biosynthesis (MONTEIRO et al., 2012), was analyzed in liquid and semi-solid medium. The production of cellulose was determined using Congo red, a dye that binds (1-4)-β-D-glucopyranosyl units (SPIERS et al., 2003). The color of the air-liquid interface ring of the M1 strain grown in semi solid medium stained with Congo red was stronger than that shown by the TRT1 ring (FIGURE 3-1A). This difference was also observed in liquid medium, where the biofilm was formed but not stained, suggesting a difference in cellulose production by these strains (FIGURE 3-1B and 1C).

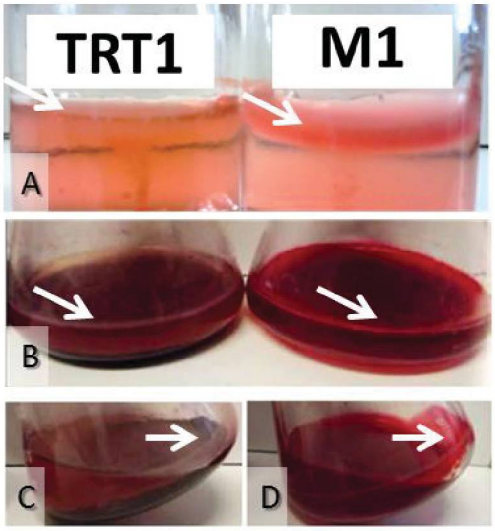
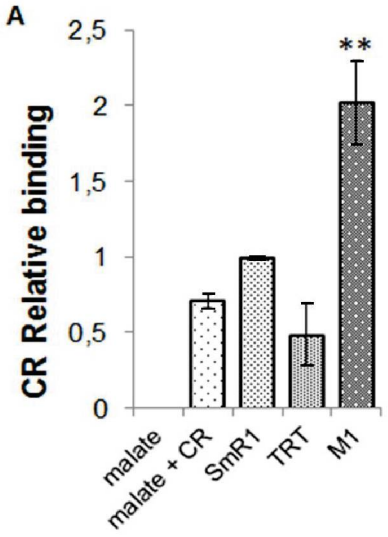


FIGURE 3-1 - Congo red staining of *H. rubrisubalbicans* strains grown in semisolid medium (A) or liquid medium (B and C) with the strains M1 and TRT1. The strains were grown in NFbHPN malate medium containing 0.005% of congo red for 3 days. The pellicule formed in the air surface (white arrows) in A and B represents a biofilm immersed in a matrix of bacterial exopolysaccharides. In C it is possible observe the no-staining biofilm adhered to the glass wall, while in D the biofilm is strongly stained by the congo red. The white arrows indicate the biofilm pellicules.

The production of cellulose was determined using Congo red. The amount of free dye in solution, after removal of the cells by centrifugation, was determined spectrophotometrically at 550 nm; the amount of bound dye was calculated by difference and was considered to be proportional to the amount of cellulose attached to the bacterial cells. The results showed that strain M1 produced more than double the amount of Congo red binding material than the TRT1 strain (FIGURE 3-2), indicating that the *bcsZ* mutant is defective in cellulose production. Ausmees et al, 1999 and Zevenhulzen et al., 1986 showed a correlation between congo red dye and cellulose content in bacterial cultures. The *bcsZ* (or *celC* in some bacteria) gene encodes for a cellulase (glycosyl hydrolase family 8), this enzyme is present in all cellulose-producing species. *Agrobacterium* CelC cellulase seems to incorporate a lipid-linked oligosaccharide intermediate into cellulose (MATTHYSSE et al.,1995). *Rhizobium* CelC2 cellulase modulates the cellulose fibrils length that mediates firm adhesion of the bacteria promoting biofilm formation (ROBLEDO et al., 2012). Mutant strains of *Rhizobium leguminosarum* bv. *trifolii* (*celA*, *celB*, *celE*, *celR2* genes) (AUSMESS et al., 1999) impaired in cellulose production also showed a reduction in congo red binding. A mutation in the *K. xylinus bcsZ* gene showed a decrease in cellulose production and an irregular packing of its fibrils was observed (NAKAI et al., 2013).



B

Status	CR Relative binding
Malate + CR	0,7088 ± 0,0477
SmR1	1
M1	1,8745 ± 0,4141
TRT	0,6119 ± 0,4257

FIGURE 3-2 - Relative Congo red binding. Total Congo Red (CR)-binding was determined by measuring mg bound-CR/OD_{600}. Relative CR-binding is calculated relative to *H. seropedicae* SmR1 – a non cellulose producer. In (A) graphs of relative binding to the CR; the controls used were only the medium malate with the dye congo red and the strain SmR1 *Herbaspirillum seropedicae* as a negative control, since this strain is unable to produce cellulose. The SmR1 strain was considered with a relative binding of one (1) to normalize the data. The assay was performed according to Spiers (2003). The results represent the average of 3 experiments performed in duplicate and there are significant differences with significance level of p> 0,01 between all analyzed points (student t test, Assistat program)

The rate of bacterial cell sedimentation might be altered by the level of exopolysachacaride production (KILLINY et al., 2013). The *bcsZ* mutant TRT1 showed a lower rate of sedimentation than the wild type strain M1 (FIGURE 3- 3). This result contrasts with that found for *Xyllela fastidiosa*, where the EPS mutant showed an increase in the sedimentation rate (KILLINY et al., 2013). The decrease in sedimentation rate shown by TRT1 was probably due to the smaller number of bacterial aggregates when compared to the large number aggregates produced by M1 under these growth conditions.

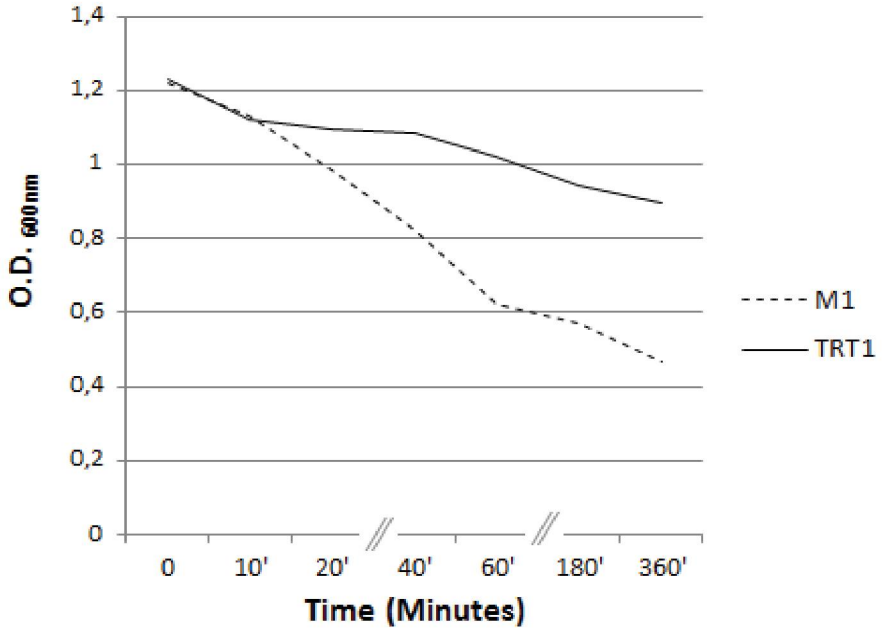


FIGURE 3-3 – Sedimentation assay. A measure of optical density in 600nm was made during six hours in the indicated times.

Influence of *H. rubrisubalbicans* M1 cellulose in biofilm formation on glass fibers.

Cellulose is present in biofilm in numerous microorganisms (AUGIMERI et al., 2015; ROMLING and GALPERIN, 2015). The presence of *H. rubrisubalbicans* M1 cellulosic material in biofilm was tested using glass fiber as supporting surfaces. The bacteria were grown in the presence of glass fibers until early visible biofilm formation. The cells were stained with calcofluor a fluorescent dye that binds β(1 → 4)-linked glucan molecules, such as chitine and cellulose, and fluoresces under UV light 360nm (HERTH et al., 1980). Biofilm formation starts with bacterial cell aggregation (FIGURE 4A), that grow three-dimensionally (FIGURES 4B and 4C). Biofilm formed on the glass fiber support was intensively stained with calcofluor (FIGURES 4D and 4F) indicating the presence of high amounts of cellulose. The presence of this polymer in the biofilm matrix of *H. rubrisubalbicans* was confirmed by the lack or low level of calcofluor fluorescence after treatment with cellulase (FIGURE 3-4E).

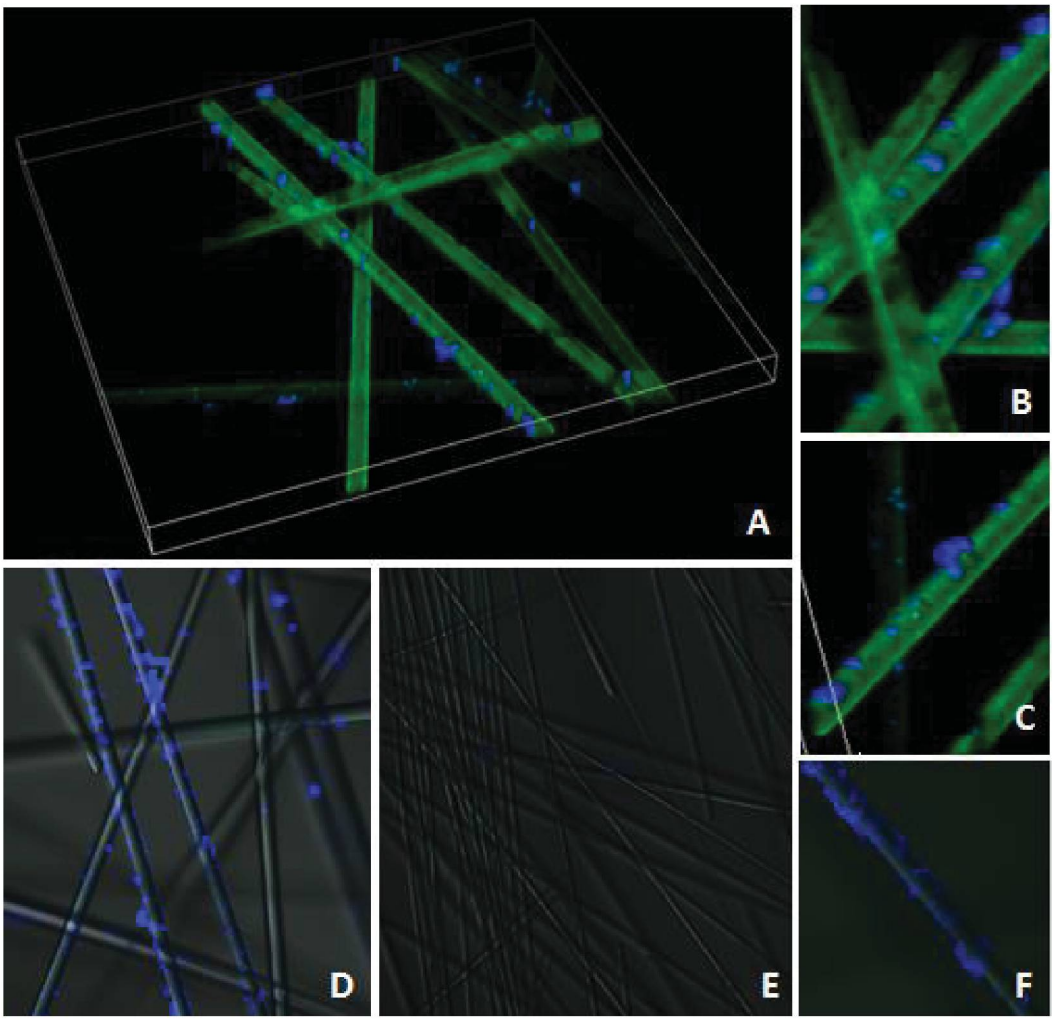


FIGURE 3-4 – Confocal microscopy of biofilm in abiotic surface. The bacteria *Herbaspirllum rubrisubalbicans* M1 was growth for 18 hours in glass wall. (A, B, C, D and F) shows the control assay, which the bacteria was incubate with only buffer for 2 hours at 30°C and thereafter was staining with calcofluor and observed in confocal microscopy. In (E), the same bacteria was growth for 18 hours in glass wall and was incubate with cellulase enzyme and buffer for 2 hours at 37°C, thereafter, the bacteria was staining and observed in confocal microscopy.

The kinetics of early biofilm formation in glass fibers by strains M1 and TRT1 of *H. rubrisubalbicans* was followed by the increase in cell numbers attached to glass fibers at 2, 4, 6 and 8 hours of growth (FIGURE 3-5). The numbers of TRT1 bacteria attached to the glass fibers were significantly smaller than those of the wild type strain throughout the experiment (FIGURE 3-5).

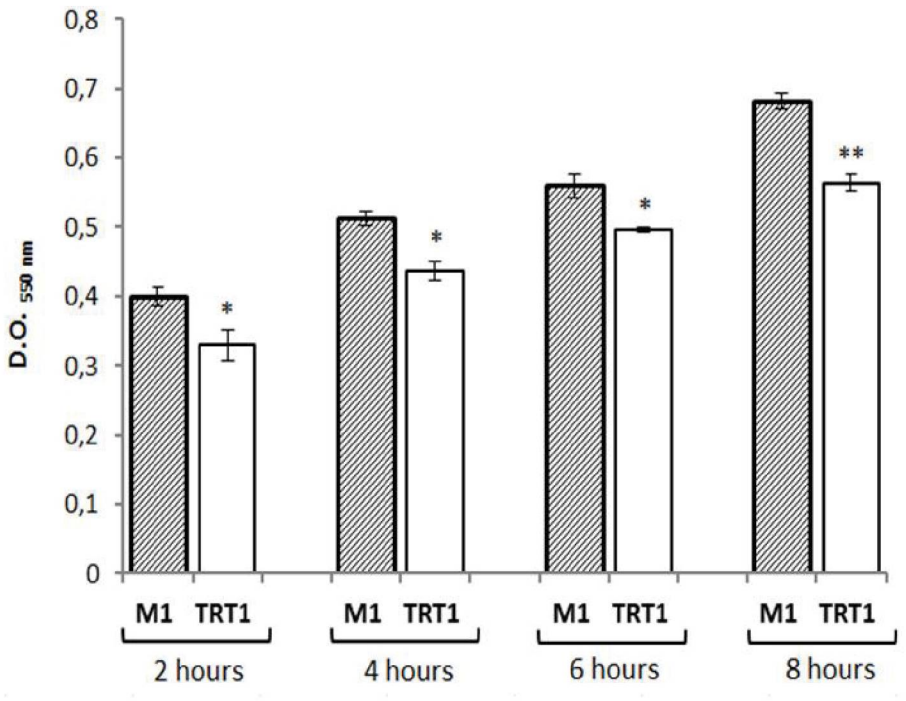


FIGURE 3-5 - Early Biofilm formation in inert matrix, by gentian violet staining, after 2, 4, 6 and 8 hours of inoculation at 30°C, 120 rpm. The assay was performed with the M1 and the TRT1 *H. rubrisubalbicans* strains. The results represent the average of three independent assays performed in duplicate. (*) Significant difference between M1 and TRT1 with a significance level of $p \leq 0.05$, (**) significant differences between M1 and TRT1 with a significance level of $p > 0.01$ (student t test, Assistat program).

Scanning electron microscopy was performed at the same time points (Figure 6). The results showed significant differences between early biofilm formation at 6 and 8 hours of growth, suggesting that the product of *bcsZ* gene is important for the formation of bacterial biofilms during static and shaken growth in liquid medium. The micrographs show that the pellicle of bacteria is more organized and compact in the wild type strain M1 when compared to those of the *bcs*$^-$ TRT1 and the fibrils is much more present in M1 aggregates (FIGURES 3-6D and 6H).

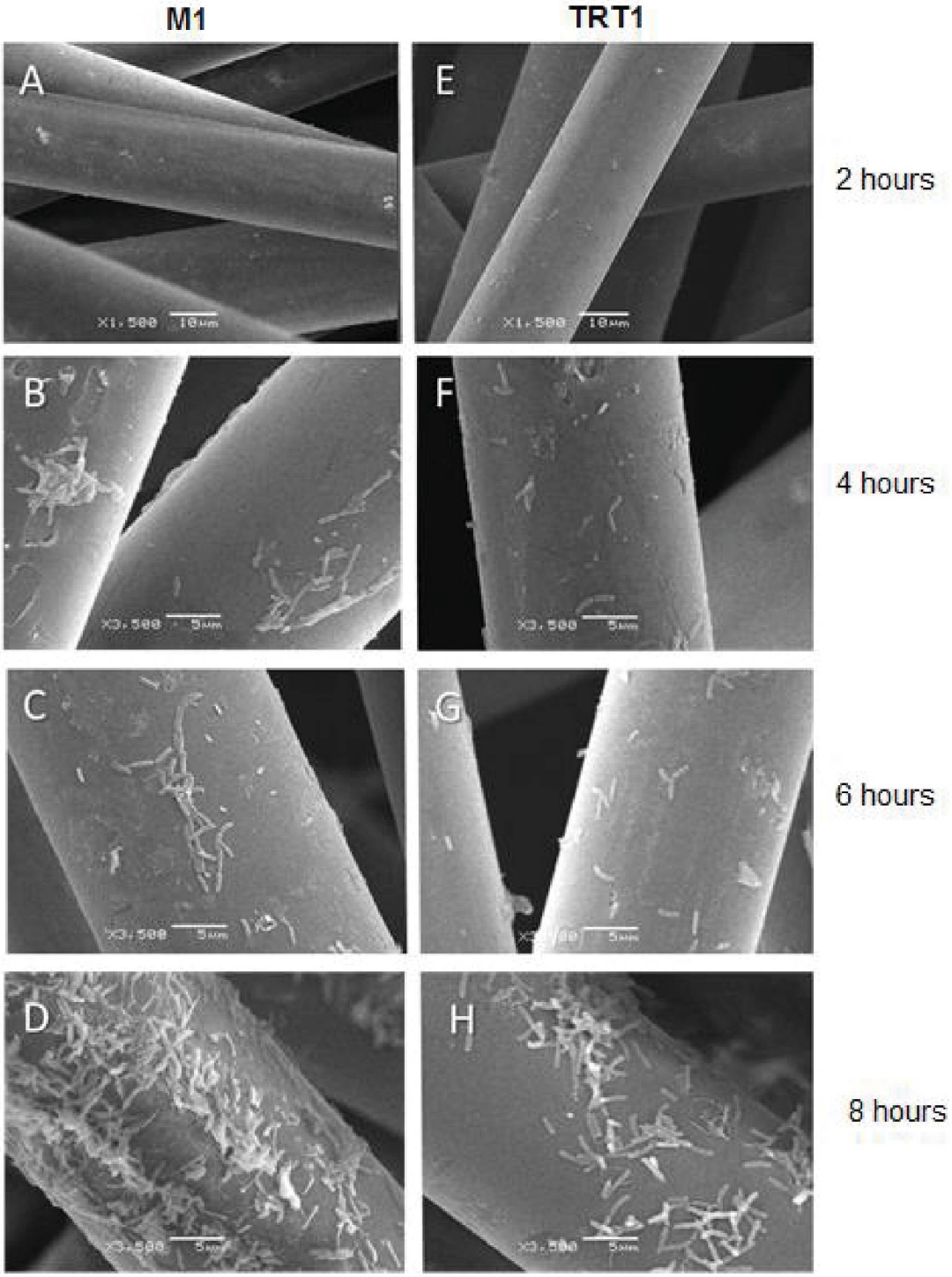


FIGURE 3-6 – Early biofilm formation on glass fibers observed by SEM. *H. rubrisubalbicans* strains were inoculated in the initial OD_{600} of 0.01 in NFbHPN malate medium under agitation (120 rpm) in the presence of glass wool, and were observed by SEM. Slides A, B, C and D shows M1 strain growth after 2, 4, 6 and 8 hours of inoculation, respectively. Slides E, F, G and H represent the TRT1 mutant growth 2, 4, 6 and 8 hours after inoculation, respectively. The increased used and scales are shown in micrographs.

The presence of cellulose-like molecules in the biofilm matrix attached to glass fibers was revealed by staining with calcofluor after 24 hours of bacterial growth. (FIGURE 3-7).The results showed a lower production of β(1-4) linked glucans by strain TRT1 as compared to the wild type M1 during biofilm formation. These data corroborate several studies regarding the importance of polysaccharides biosynthesis in the bacterial aggregation/biofilm formation processes. Furthermore, the data confirmed that strain TRT1 is unable to produce cellulose and that cellulose is involved in biofilm formation in the wild type strain M1 of *H.rubrisubalbicans*.

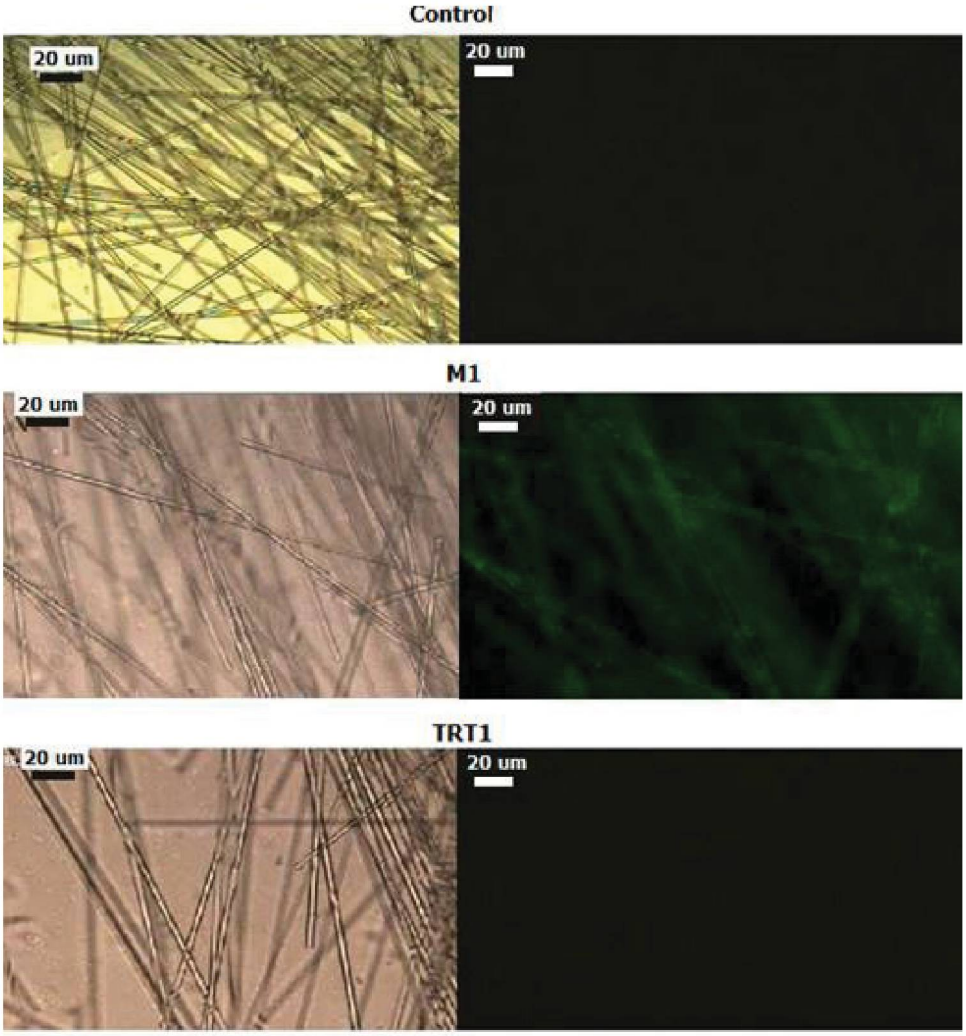


FIGURE 3-7 - *H. rubrisubalbicans* biofilm formation on glass fibers stained with calcofluor. Fluorescence microscopy assay on glass fiber dyed with calcofluor showing the cellulose presence only in M1 *H.rubrisubalbicans* strain. The left images shows light microscopy and the right images show the same region under fluorescence microscopy. Samples were observed under a microscope Carl Zeiss - Jena, with a 25x objective, images were captured by AxioCam camera and AxioVision software MRc LE 4.6 was used to obtain the images. These pictures are representative of three independent experiments performed under the same conditions.

H. rubrisubalbicans formed biofilm on glass cover slips in static liquid cultures. After 72 hours the early biofilm attached to the cover slips was analysed by Scanning Electron Microscopy (SEM) in order to determine possible morphological changes in the structural pattern of the biofilms formed by strains M1 and TRT1 of *H.rubrisubalbicans* (FIGURE 3-8).

The results showed that biofilms formed by strain TRT1 was fragile and easily removed, while those formed by strain M1 were more compact and resistant to detachment (FIGURES 3-8A, 3- 8B, 3-8C and 3-8D). As revealed by SEM strain M1 showed larger bacterial aggregate, while TRT1 had smaller and more widely spaced clusters of bacteria. Fibrils characteristic of polysaccharide fibers were observed prominently in the more organized biofilms of the wild type strain.

Given these results, we conclude that the mutation in the *bcsZ* gene leads to alterations in the bacterial early biofilm formation, changing the pattern of fibrils, as indicated by arrows in the Figures 8B and 8D. These changes lead to formation of smaller bacterial aggregates and consequent structural alteration in biofilm. Cellulose is present in bacterial biofilm from *Gluconacetobacter xylinum* (CANNON and ANDERSON, 1991), *Escherichia coli*, *Salmonella enteriditis*, *Salmonella typhimurium* (WHITE et al., 2003; ZOGAJ et al., 2001), *Pseudomonas fluorescens* (SPIERS et al., 2003). *Rhizobium sp.*, mutant in *celC2* gene, had longer cellulose microfibrils and present drastic damages in the biofilm formation (ROBLEDO et al., 2012). Mutants of *Agrobacterium tumefaciens, P. fluorescens and Salmonella enteriditis* impaired in cellulose biosynthesis have been shown to be defective in biofilm formation (MATTHYSSE et al., 2005, SPIERS et al., 2003, WHITE et al., 2003).

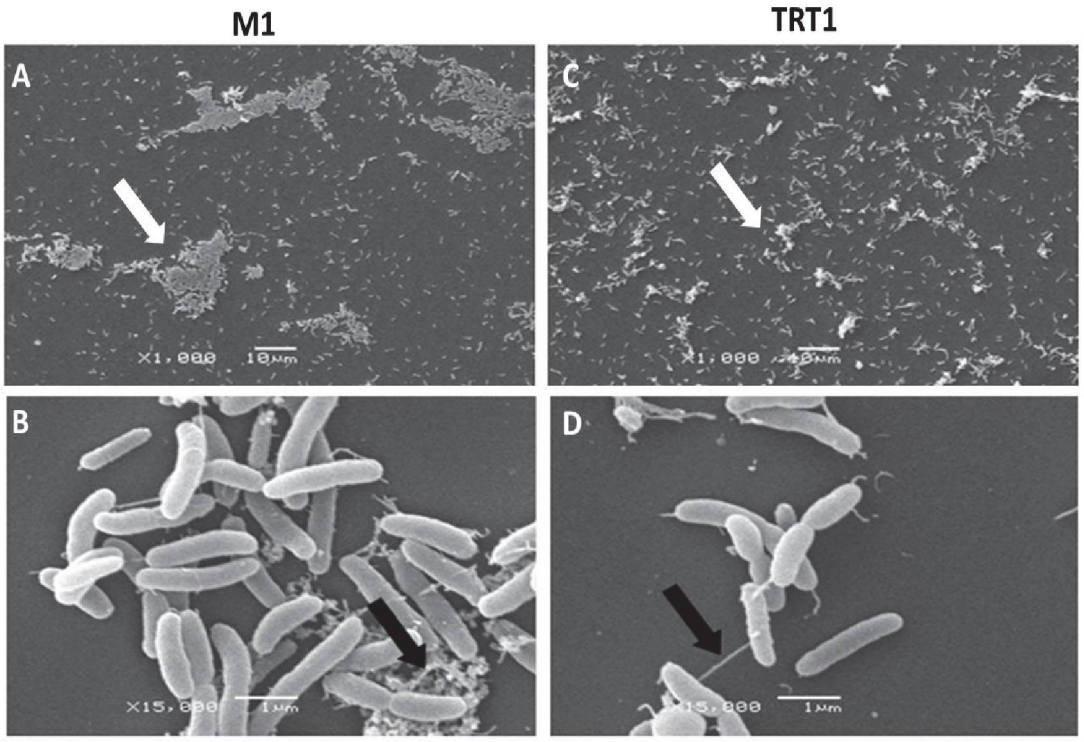


FIGURE 3-8 - Scanning Electron Microscopy of *H.rubrisubalbicans* strains M1 and TRT1. The bacteria were grown in petri dishes containing liquid medium and coverslips with 1 cm of diameter. The coverslips recovered after 72 hours were observed and the images show bacterial early biofilm formed. The images were taken on a scanning electron microscope JEOL JSM-6360 LV. The increased use and scales are shown in micrographs. A and B showed the M1 strain. C and D showed the TRT1 strain. The white arrows indicate the bacterial aggregates. The black arrows indicate the difference in the arrangement of fibrils observed in bacterial aggregates.

Influence of cellulose in *H. rubrisubalbicans* motility

The effect of cellulose level on the motility of the *bcsZ* gene mutant TRT1 was determined in semi-solid medium. Strain TRT1 was much more slower than the wild type strain, suggesting that a deficiency in cellulose production alters the motility of *H. seropedicae* as well as biofilm formation (FIGURE 3-9).

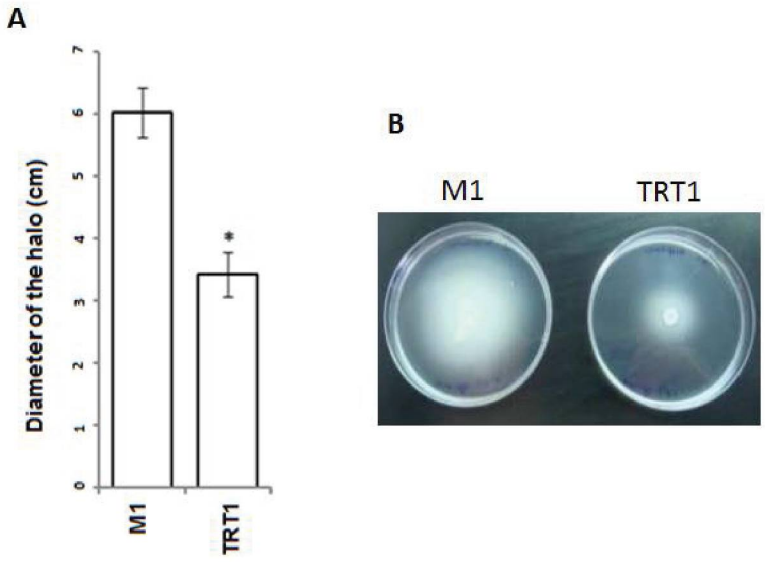


FIGURE 3-9 - Motility of *H. rubrisubalbicans* strains M1 and TRT1. The assay was performed by measuring the halo formed after growth in NFbHPN malate medium containing 0.25% agar. (A) shows the determination of the diameter of the motility halo produced by strains M1 and TRT1 and represent the average of four independent experiments. (B) shows a representative picture of the obtained data. The data show statistically significant differences between the motility of the strains at the level of p> 0.01 (t test, Assistat program).

Involvement of Cellulose in plant-*Herbaspirillum rubrisubalbicans* interaction

H. rubrisubalbicans TRT1 mutant attached to maize root surface was 53 fold lower than the wild type (MONTEIRO et al, 2012), indicating that cellulose production was involved in the colonization of maize roots.

SEM of maize root surface colonized by *H. rubrisubalbicans* strains M1 and TRT1, showed a lower formation of biofilm by TRT1 as compared to M1. To demonstrate possible alterations during epiphytic maize colonization by *H. rubrisulbalbicans* M1 and TRT1, we performed a Scanning Electron Microscopy (SEM) of the time-course of colonization 1, 3 and 7 days after inoculation. The extent of Maize roots colonization by strain M1 (FIGURES 3-10 A-C) and by strain TRT1 (FIGURES 3-10 D-F) are shown. One day after inoculation (d.a.i.) of strain M1, single cells and discrete bacteria aggregates associated to the periclinal cell wall surface by apolar attachment are seen in roots seedlings inoculated with the wild type strain (FIGURE 3-10A). A similar pattern of colonization is observed for the mutant strain TRT1 (FIGURE 3-10D). The colonization with M1 strain reached a maximum aggregates formation

7 d.a.i (FIGURE 3-10C). In contrast, the mutant strain remained as single cells with minimum small aggregate formation throughout the experiment (FIGURES 3-10D, 3-10E and 3-10F).

Treatment of *H. rubrisubalbicans* with cellulase before inoculation of maize lead to a statistically significant decrease of 6 times in root attached cells from $3x10^6$ to $5x10^5$ CFU (FIGURE 3-11). This confirms the involvement of cellulose in *H. seropedicae* colonization as previously reported (MONTEIRO et al., 2012).

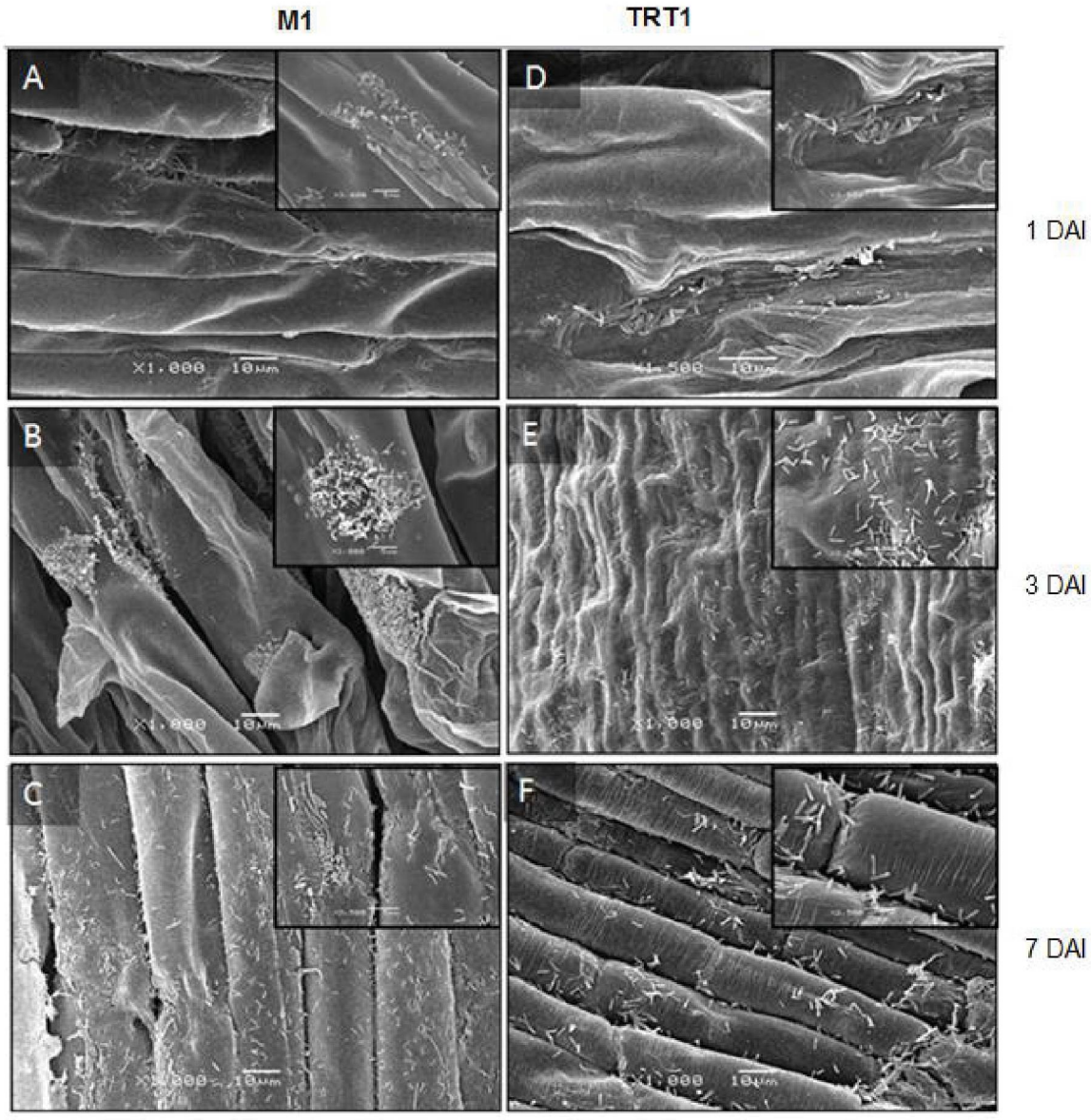


FIGURE 3-10 – Epiphytic maize colonization by *H. rubrisulbalbicans* M1 and TRT1. Maize seedlings were inoculated with 10^5 bacteria and grown in Plant medium. After 1, 3 and 7 inoculation days, the roots were prepared and analyzed by SEM. Slides A, B and C show the growth of the wild type strain M1 after 1, 3 and 7 days, respectively. Slides D, E and F represent the growth of the mutant strain TRT1 after 1, 3 and 7 days of inoculation, respectively. The images were taken on a scanning electron microscope JEOL JSM-6360 LV. The amplification and scales are shown.

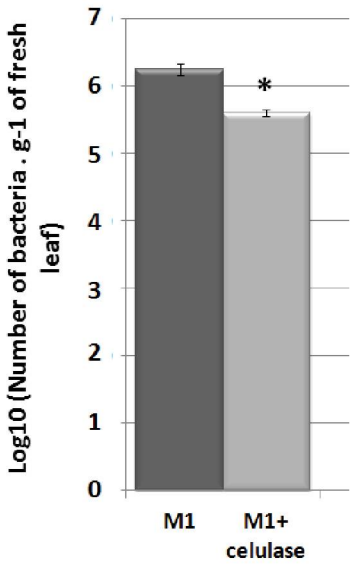


FIGURE 3-11 Number of bacteria attached to maize root, 30 minutes after inoculation, in the presence of cellulase. The bacteria were treated with 250 µg/mL of the cellulase enzyme for 2 hours and 30°C. After the treatment, the cells were washed and inoculated in maize seedlings at the concentration of $\mathbf{10^5}$ bacteria/seedlings for 30 minutes. Controls were submitted to the same procedure, and only the cellulase enzyme buffer was added. After the inoculation, the attached bacteria were recovered and counted. (*) Indicates a significant difference between treated with cellulase and untreated cells with significance level of p> 0.01 (Test T, Assistat program).

The pattern of endophytic colonization of maize roots by TRT1 and M1 was compared. Although deficient in cellulose production the *bcsZ* mutant colonized root maize endophytically, though at a lower level than the wild type strain (FIGURE 3-12). The results suggest that cellulose production by *H. rubrisubalbicans* is not essential for the colonization *of* maize roots, and other factor such as LPS and EPS could be also involved. *Agrobacterium tumefaciens* C58 *chvB* and *celA* mutant strains, deficient in cellulose production, showed a reduction in root colonization when compared with the wild type (MATHYSSE et al., 2005). In *Rhizobium leguminosarum* and *Rhizobium trifolii* cellulose fibrils are important in the bacterial attachment to *Vicia sativa* root hairs (LAUS et al., 2005, DAZZO et al., 1984). The capacity of *Rhizobium leguminosarum* bv. *trifolii celC2* mutants to form biofilm on plant roots was also significantly reduced (ROBLEDO et al, 2012).

Ours results indicate that the production of cellulose is involved in *H. rubrisubalbicans* cell aggregation, biofilm formation and maize root colonization.

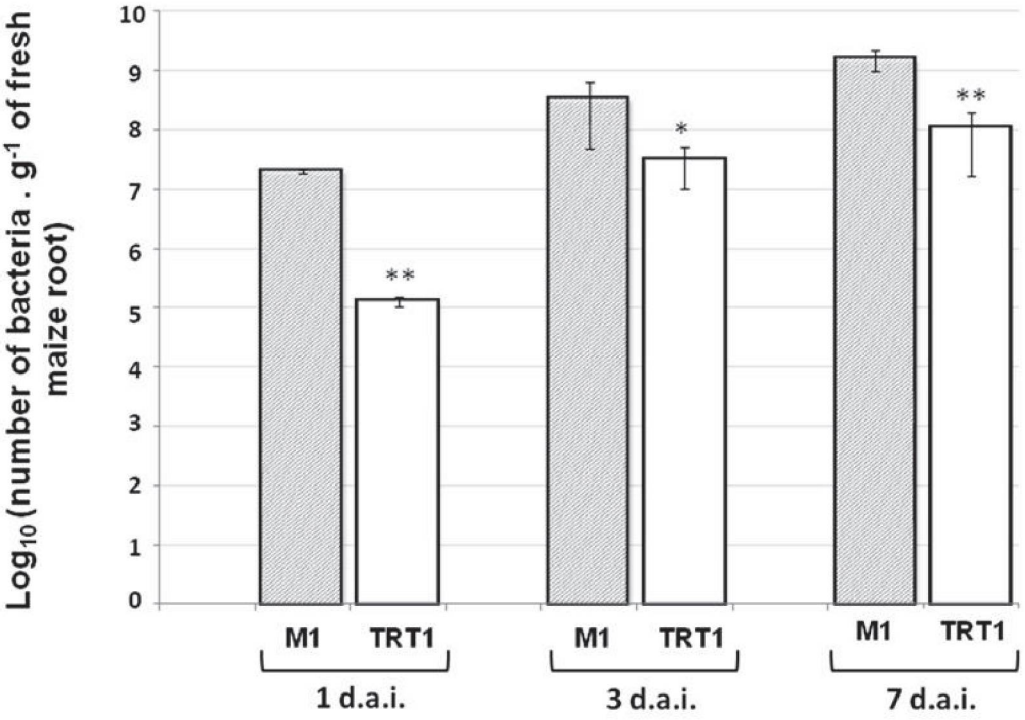


FIGURE 3-12 - Endophytic maize root colonization by *H. rubrisubalbicans* M1 (grey bars) and TRT1 (white bars) a bcsZ mutant strain. The numbers of endophytic cells were also determined 1, 3 and 7 days after inoculation (d.a.i.). The data represent the average of at least 3 biological replicata with 5 experimental determination each. (*) Significant statistical differences were observed between M1 and TRT1 with a significance level of p> 0:05. Significant difference between M1 and TRT1 with a significance level of p≤0.05(*) and p≤0.01 (**) (t test, Assistat program)

MATERIAL AND METHODS

Strains and growth conditions

The strains of *H. rubrisubalbicans* M1 (BALDANI et al., 1996) and TRT1 (MONTEIRO, et. al. 2012) were grown in NFbHP malate media (KLASSEN et al., 1997) with 20 mmol/L of NH_4Cl as nitrogen source. The cultures were incubated under agitation (120 rpm) for 18 – 24 hours, at 30°C. Semi-solid and solid NFbHP malate media contained 0,175 and 1,5% (w/v) agar, respectively. To medium destined to the growth of the TRT1 strain contained kanamycin at a final concentration of 500 µg/mL.

Effect of cellulase on *H. rubrisubalbicans* attachment to maize roots

The M1 strain was grown up to $O.D._{600}$ = 1,0, treated with 250 µg/mL cellulase (Sigma Aldrich) for 2 hours at 30°C and washed 3 times with sterile saline solution (NaCl 0.9%). The

treated bacteria were inoculated to maize seedling 2 days old at a concentration of 10^5 bacteria/seedling. After 30 minutes exposure the seedlings were washed 3 times with sterile saline solution, and the root attached bacteria were recovered by vortexing for 45 seconds in sterile saline solution, the seedlings were removed and the bacteria present in the supernatant suspension was counted by serial dilution.

Motility assay

Aliquots (20µL) of bacterial suspensions at the concentration of 10^8 bacteria.mL^{-1} were spot inoculated in NFbHP malate medium containing 0,25% agar in petri dishes. The cultures were incubated at 30°C without agitation, and the diameter of growth halo was determined after 48 hours.

Sedimentation Rate

The bacterial strains were grown in liquid NFbHPN malate media at 120 rpm, and 30°C, up to $O.D._{600}$ =1. After that the culture was kept static and the supernatant was measured for the following times: 10, 20, 40, 60, 180 and 360 minutes.

Biofilm formation in the inert matrix of glass fibers/wool

Biofilm formation by *Herbaspirillum sp*. strains was carried out in liquid media in the presence of 50 mg of glass wool. The cultures were incubated at 30°C, 120 rpm for the specified time. After incubation, the glass wools was recovered, washed with 0.9% NaCl solution and stained with 200 µL of gentian violet for 5 minutes. The glass wools were washed five times with NaCl 0,9% solution to remove the gentian violet non-binding to the sample, and finally washed with 1 mL of 70% ethanol to decolorize the glass wool and the absorbance at 550 nm measured in order to determine the presence of biofilm.

To visualize the pattern of biofilm formation, the bacteria were grown on petri plates containing NFbHPN malate liquid and glass slides of approximately 1 cm in diameter on the bottom. After 72 hours of static growth at 30° C, the coverslips were recovered, fixed and prepared for SEM. Similarly, the glass wools were also recovered after the bacteria growth, prepared and observed at SEM, as described below.

The samples were fixed in Karnovisky solution (2% paraformaldehyde, 2.5% glutaraldehyde in 0.1M cacodylate buffer pH 7.2 at 4° C) (KARNOVSKY et al., 1965), dehydrated in ascending alcohol and acetone series, and the critical point was reached in a

Bal-Tec CPD - 030 with carbon dioxide. The plating in gold was obtained in a Balzers SCD - 030. SEM analysis were performed in a scanning electron microscope JEOL-JSM 6360 LV.

Staining of *H. rubrisubalbicans* cells with congo-red

Congo red staining was made by adding a 0.005% Congo red solution to liquid or semi-solid NFbHP-malate static cultures of *Herbaspirillum* spp. after 2 days of growth, when biofilm formation occurs. The stained biofilm was recovered, vortexed, centrifuged and the optical density of supernatant was determined spectrophotometrically at 550 nm. Bound dye was calculated by difference.

Since *Herbaspirillum seropedicae* strain SmR1 lacks all cellulose biosynthesis genes and does not produce cellulose (MONTEIRO et al., 2012) it was used as reference to determine the binding of congo red relative to the *H. rubrisubalbicans* strains.

Staining of *H. rubrisubalbicans* cells with calcofluor

Calcofluor staining was carried out in *H. rubrisubalbicans* biofilm produced on an inert matrix. The strains were grown in liquid NFbHPN malate medium in the presence of 50 mg of glass wool at 30°C and 120 rpm for 24 hours. The glass wool was recovered and stained with 50 mM calcofluor for 1 hour, washed with NaCl 0.9% and analyzed in a confocal microscope.

To determine the role of cellulose on biofilm stability, it was first treated with cellulase (250 µg/mL), then stained with calcofluor and observed in a confocal microscope. The samples were analyzed by visible light and fluorescence (360nm excitation) microscopy with a Carl Zeiss Jena microscope; the images were obtained using an AxioCam camera attached to the microscope. The AxioVision LE software 4.6. was used to treat the images. Confocal images were obtained on a Nikon Ti Microscope. Tridimensional images were obtained with NIS-Elements software – Nikon.

Maize – *Herbaspirillum rubrisubalbicans* interactions

Two maize cultivars SHS 3031 and SHS 5055 were used in this study. Maize seeds were sterilized for 20 minutes in a solution containing 1% sodium hypochlorite and 0.4% tween 20%. The seed were then washed with 70% ethanol for 5 minutes and 4 times with sterile MilliQ water for 1 minute. Seeds were germinated in 1% agar-water plates for 48-72 hours at 28 - 30°C in the dark. The seedlings were inoculated with bacterial suspensions containing 10^5 bacteria/mL for 30 minutes. The seedlings were then transferred to

experimental tubes containing 20 mL of Plant Medium (EGENER et al., 1999) solution and polypropylene spheres and were kept at 25°C with a 12 hours of photoperiod cycle. To determine the total number of root attached bacteria, the plants were recovered, washed 3 times with a saline (NaCl 0,9%) solution, macerated, serially diluted and plated on NFbHPN malate medium. To visualize biofilm formation by SEM, the roots were collected, fixed in Karnovisky solution and prepared in the same way as the samples described above. The samples were analyzed in a Scanning electron microscope JEOL-JSM 6360 LV.

CONCLUSION

The mutation in the gene *bcsZ* affected EPS production, as shown in mutant strain TRT1 of *H.rubrisubalbicans*, suggesting that occurred a reduction of the synthesis of cellulose. The *bcsZ* gene mutation altered motility and biofilm production of mutant strain TRT1, indicating that the cellulose biosynthesis may be important for both these processes.

A mutation in the *wss* operon, although it is important to bacteria plant interaction, as shown in the colonization assays in maize roots. These results show that the biosynthesis of cellulose is important for the interaction between *H.rubrisubalbicans* and maize, but mainly important to the biofilm production.

AUTHOR CONTRIBUTION

Lucélia Donnatti (Department of Cellular Biology – UFPR) – Microscopic analysis
Fabio de Oliveira Pedrosa, Emanuel Mantempi de Souza and Rose Adele Monteiro (Department of Biochemetry and Molecular Biology) – Orientation and Supervision.

ACKNOWLEDGMENTS

We are grateful to Roseli Prado and Valter A. Baura for their technical support and to Electron Microscope Center - Federal University of Parana for the microscope images. This work was supported by CNPq/INCT, CAPES and Fundação Araucária.

REFERENCES

ASSMUS, B.; HUTZLER, P.; KIRCHHOF, G.; AMANN, R.; LAWRENCE J. R.; HARTMANN, A.; In Situ Localization of Azospirillum brasilense in the Rhizosphere of Wheat with Fluorescently Labeled, rRNA-Targeted Oligonucleotide Probes and Scanning Confocal Laser Microscopy. **Applied and Environmental Microbiology**. vol. 61 no. 3 1013-1019. 1995.

AUSMEES, N., JONSSON, H., HOGLUND, S., LJUNGGREN, H. AND LINDBERG, M. Structural and putative regulatory genes involved in cellulose synthesis in Rhizobium leguminosarum bv. trifolii. **Microbiology** 145:1253-1262. 1999.

BALDANI, J.I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum* seropedicae gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. **International Journal of Systematic Bacteriology**. V. 36. P. 86-93. 1986.

BALDANI, V.L.D.; BALDANI, J.I.; OLIVARES, F.; DÖBEREINER, J. Identification and ecology of *Herbaspirillum* seropedicae and the closely related Pseudomonas rubrisubalbicans. **Symbiosis**, v. 13, p. 65-73, 1992.

BALDANI B., POT G., KIRCHHOF E., FALSEN V. L. D., BALDANI F. L., OLIVARES, B. HOSTE, K. KERSTERS, A. HARTMANN, M. GILLIS, J. DÖBEREINER. Emended Description of *Herbaspirillum*; Inclusion of [Pseudomonas] rubrisubalbicans, a Mild Plant Pathogen, as *Herbaspirillum* rubrisubalbicans comb. nov.; and Classification of a Group of Clinical Isolates (EF Group 1) as *Herbaspirillum* Species 3. **Int J Syst Bacteriol**, v. 46, p.802- 810, 1996.

BALDANI, V.L.D.; BALDANI, J.I.; DOBEREINER, J. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum* seropedicae and Burkholderia spp. **Biol. Fertil. Soils** 30:485-491, 2000.

EGENER, T., HUREK, T., AND REINHOLD-HUREK, B. Endophytic expression of nif genes of Azoarcus sp. strain BH72 in rice roots. **Mol Plant Microbe Interact** 12: 813–819. 1999.

FUQUA, T. and CLAY, T. Biofilm Formation by Plant-Associated Bacteria. **Annu. Rev. Microbiol.** 61:401–22. 2007.

GORIN P. A. J., TEIXEIRA, A. Z. A., TRAVASSOS LR, LABOURIAU MLS, IACOMINI M. Characterization of carbohydrate omponents of an unusual hydrogel formed by seed coats of Magonia pubescens(Tingui). **Carbohydrate Research** 282: 325 – 333. 1996.

HARSHEY RM, MATSUYAMA T. Dimorphic transition in Escherichia coli and Salmonella typhimurium: surface-induced differentiation into hyperflagellate swarmer cells. **Proc Natl Acad Sci U S A**. 1994 Aug 30;91(18):8631-5.

HENRICHSEN, J.; Bacterial Surface translocation: a survey and cassification. **Bacteriol Rev.** 36: 478-503. 1972.

HERTH, W., SCHNEPF. E. The Fluorochrome, Calcofluor White, Binds Oriented to Structural Polysaccharide Fibrils.**Protoplasma** 105, 129-133 (1980)

JAHN, A. et al. Composition of Pseudomonas putida biofilms: accumulation of protein in the biofilm matrix. **Biofouling** 14, 49–57. 2000.

JAMES, E. K.; OLIVARES, F. L.; BALDANI, J. I.; DÖBEREINER, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of Sorghum bicolor L. Moench. **Journal of Experimental Botany.** v. 48. n. 308. p. 785-797. 1997.

KANG, K. S; COTTRELL, I. W. Polysaccharides in Microbial Technology. **Academic Press: New York**, 1979, p. 417

KARNOVSKY, M. J. A formaldehyde-glutaraldehyde fixative of high osmolality for use in electron microscopy. **J Cell. Biol**. 27: 137-138. 1965

KASANA, R. C., SALWAN, r., DHAR, H., DUTT, S., GULATI, A., A Rapid and Easy Method for the Detection of Microbial Cellulases on Agar Plates Using Gram's Iodine. **Current Microbiology**. V.57 (503-507). 2008.

KILLINY, N.; MARTINEZ, R. H.; DUMENYO, C.K.; COOKSEY, D. A.; ALMEIDA, R. P. The exopolysaccharide of Xylella fastidiosa is essential for biofilm formation, plant virulence, and vector transmission. **Molecular Plant Microbe Interaction**. 2013 Sep;26(9):1044-53

KLASSEN, G.; PEDROSA, F. O.; SOUZA, E. M.; FUNAYAMA, S.; RIGO, L. U. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum* seropedicae SMR1. **Can. J. Microbiol**. 43:887-891, 1997.

KOLTER R., GREENBERG, E. P., News & views feature The superficial life of microbes. **Nature**, 2006.

MATTHYSSE, A.G., HOLMES, K.V. AND GURLITZ, R.H. Elaboration of cellulose fibrils by Agrobacterium tumefaciens during attachment to carrot cells. **J Bacteriol** 145:583-595. 1981.

MONTEIRO, R. A., BALSANELLI, E.; TULESKI, T. R.; FAORO, H.; CRUZ, L. M.; WASSEM, R.; BAURA, V. A,; TADRA-SFEIR, M. Z.; WEISS, V.; DAROCHA, W. D.; MULLER-SANTOS, M.; CHUBATSU, L. S.; HUERGO, L. F.; PEDROSA, F. O. AND SOUZA, E. M. Genomic comparison of the endophyte *Herbaspirillum* seropedicae SmR1 and the phytopathogen *Herbaspirillum* rubrisubalbicans M1 by Suppressive Subtractive Hybridization and partial genome sequencing. **FEMS Microbiology Ecology**, 2012.

OLIVARES F. L.; JAMES E. K.; BALDANI J. I, DOBEREINER J. Infection of mottled stripe disease-susceptible and resistant sugar cane varieties by the endophytic diazotroph *Herbaspirillum*. **New Phytol**. vol.135, p. 723-737, 1997.

OTTEMANN, K. M.; MILLER, J. F.; Roles for motility in bacterial-host interactions. Mol **Microbiol.** 1997 Jun;24(6):1109-17.

PEHL, M.J.; JAMIESON W. D.; KONG, K.; FORBESTER J. L.; FREDENDALL, R. J.; GREGORY, G. A.; MCFARLAND, J. E.; HEALY J. M.; ORWIN, P. M.; Genes that influence swarming motility and biofilm formation in Variovorax paradoxus EPS. **Plos One.** 2012;7(2):e31832

RASHID MH, KORNBERG A. Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa. **Proc Natl Acad Sci U S A.** 2000 Apr 25;97(9):4885-90.

ROBLEDO, M; RIVERA, L; JIMÉNEZ-ZURDO, JI; RIVAS, R; DAZZO, F; VELÁZQUEZ, E; MARTÍNEZ-MOLINA, E; HIRSCH, AM; MATEOS, P. F. Role of Rhizobium endoglucanase CelC2 in cellulose biosynthesis and biofilm formation on plant roots and abiotic surfaces. **Microbial Cell Factories** 11:125, 2012.

ROMLING, U. Molecular biology of cellulose production in bacteria. **Res Microbiol** 153: 205-212. 2002.

RONCATO-MACCARI, L. D. B., RAMOS, H. J.O., PEDROSA, F. O., ALQUINI, Y., CHUBATSU, L. S., YATES, M.G., RIGO, L.U.,STEFFENS, M.B., SOUZA, E. M. Endophytic *Herbaspirillum* seropedicae expresses nif genes in gramineous plants**. FEMS Microbiology ecology**, 45(37-47), 2003

SMIT, G. KIJINE, J.W. LUGTCNBERG, B .J . J. Involvement of both cellulose fibrils and a Ca2+ -dependent adhesin in the auachement of Rhizobium leguminosarum to pea root hair tips. **Bacteriol.**1 69:4294-430. 1987.

SOLANO, C.; GARCÍA, B.; VALLE, J., BERASAIN, C.; GHIGO, J.; GAMAZO, C.; LASA, I.; Genetic analysis of Salmonella enteritidis biofilm formation: critical role of cellulose. **Molecular Microbiology**. 43(3), 793-808; 2002.

SPIERS, A. J., BOHANNON, J., GEHRIG, S. AND RAINEY, P.B. Colonization of the air-liquid interface by the Pseudomonas fluorescens SBW25 wrinkly spreader requires an acetylated form of cellulose. **Molecular Microbiology** 50: 15–27. 2003

VERSTRAETEN, N., BRAEKEN, K., DE BKUMARI, B., FAUVART, M., FRANSAER, J. Living on a surface: swarming and biofilm formation. **Trends in microbiology**, 16: 496-506. 2008

WALKER T.S.; BAIS, H.P.; DEZIEL, E.; SCHWEIZER, H.P.; RAHME, L.G.; Pseudomonas aeruginosa-plant root interactions. Pathogenicity, biofilm formation, and root exudation. **Plant Physiology;** 134:320–31. 2004

WALL D.; KAISER, D. Type IV pili and cell motility. **Molecular Microbiology**. 1999 Apr;32(1):1-10.

WATNICK, P. And KOLTER, R. MINIREVIEW - Biofilm, City of Microbes. **Journal of bacteriology**., p. 2675–2679 Vol. 182, No. 10. 2000.

WOLFROM, M.; THOMPSON, A. Acetylation. **Methods in carbohydrate chemistry**, v. 2, n. 2, p. 211-215, 1963a.

WOLFROM, M.; THOMPSON, A. Reduction with sodium borohydride. **Methods in carbohydrate chemistry**, v. 2, n. 1, p. 65-67, 1963b.

CAPÍTULO IV

Manuscrito ainda não submetido

Herbaspirillum rubrisubalbicans M1 AS A PATHOGENIC MODEL TO STUDY THE IMMUNE SYSTEM OF *SORGHUM BICOLOR*

Herbaspirillum rubrisubalbicans as a pathogenic model to study the immune system of *Sorghum bicolor*

ABSTRACT

Herbaspirillum rubrisubalbicans is a mild pathogen of some varieties of sugar cane and sorghum. In Sorghum bicolor this bacteria causes the red stripe disease and in sugar cane it is responsible for the mottle stripe disease. In Sorghum it affect leaves leading to tissue necrosis, reduction in photosynthesis and impaired leaf development. In this work, fifty-eight genotypes of sorghum were inoculated with *H. rubrisubalbicans* and a total of 54 of those showed disease symptoms after 7 days. Bacteria were recovered from portions of the leaf increasingly distal from the site of inoculation. The results showed the spread of both bacteria and symptoms along the expanding leaf. Plants have the ability to mount a defense response to invading pathogens due to recognition of conserved motifs, termed microbe-associated molecular patterns (MAMPs). We tested the ability of MAMP-triggered immunity to protect sorghum plants from *H. rubrisubalbicans* infection. Sorghum leaves pre-treated with chitin and flg22 (a conserved 22 amino acid peptide derived from bacterial flagellin) and then challenged with the pathogen were protected. RNAseq and QTL analysis of these plants revealed classic pathways of plant pathogen interaction, suggesting that *H. rubrisubalbicans* is a useful pathosystem to study innate immunity response of sorghum.

INTRODUCTION

Sorghum bicolor is the fifth most widely grown crop in the world, it is an important grass with several features that make it a great candidate for bioenergy production (ROONEY et al., 2007). Among the important features are the yield potential and composition, the already established production system, the rapid growth characteristics, the high carbon content and the potential for genetic and breeder improvements. Another crucial feature is that the sorghum production can be associated with hot and dry environments and shows a high water-use efficiency and drought tolerance (BOYER et al., 1982, WIEDENFELD et al., 1984 and ROONEY et al., 2007). Combined with all this characteristics the energy crisis increase the interest in explores the development of a sweet sorghum for bioenergy, since it has strong potential as a renewable energy resource (biofuels and energy production) (MCBEE et al., 1987).

Different microbes can be found in beneficial or pathogenic association with *Sorghum bicolor,* such as *Herbaspirillum*. The *Herbaspirillum* genus comprises a gram-negative β-Proteobacteria group, the most part of the species are plant associative diazotrophic bacteria, being found in a beneficial association: increasing the growth and productivity of several crops such as maize (*Zea mays*), sugar cane (*Saccharum officinarum*) and rice (*Oryza sativa*) (BODDEY et al. 1995; JAMES 2000; JAMES et al., 2002; ELBELTAGY et al., 2001; GYANESHWAR et al., 2002; RONCATO-MACCARI et al., 2003).

One of the *Herbaspirillum* species already described is *Herbaspirillum rubrisubalbicans* which is also a diazotrophic organism and can be isolated from leaves and roots of healthy grasses (BALDANI et al. 1992; DOBEREINER et al. 1994; OLIVARES et al. 1996; CRUZ et al. 2001). In contrast, this bacterium was also isolated from sick sugar cane leaves and has been described as responsible for the mottle stripe disease in some sensitive genotypes of sugar cane and the red stripe disease in susceptible genotypes of *Sorghum bicolor* (PIMENTEL et al., 1991; BALDANI et al., 1992, HALE and WIKIE, 1972 and PIMENTEL et al., 1991). These diseases can reduce 50% the lifetime of the leaves but does not kill the plant and even under pathogenic behavior the bacterium is capable of express the nitrogenase genes and fix the atmospheric nitrogen (JAMES et al. 1997; OLIVARES et al. 1997). In 1980 Galli and collaborators reported a susceptible cultivar (B-4362) in Barbados and in 2010 the mottle stripe disease was also described in sugar cane in China (Tan, 2010). However, according to Pimentel (1991) and Olivares (1997) there is no report of mottled/red stripe disease spread in the Brazilian sugarcane and sorghum fields, being all the agronomically important genotypes resistant for this disease. Olivares, 1997 also showed that the artificial inoculation in a widely planted sugarcane cultivar (cv. SP 70-1143) in Brazil in those years did not show any symptom of the disease after inoculation.

Herbaspirillum rubrisubalbicans can be found colonizing the internal tissues of the host such as substomatal cavities, intercellular spaces and xylem (OLIVARES et al., 1995), the bacteria can also be found in different areas of the leaf (far from the inoculation point) suggesting a free translocation of the bacteria in the host, probably via the vascular vessels (PIMENTEL et al., 1991; BALDANI et al., 1992; DOBEREINER et al., 1993). The interaction between *Herbaspirilum rubrisubalbicans* and *Sorghum* is partially unknown, especially by the plant side. Being the *Sorghum bicolor* a grass with different characteristics that support it as an important bioenergy crop it is important

understand the mechanism behind this interaction. It is also very important to develop a model for the study of the immune system immune system of Sorghum bicolor, since this plant has a great potential to face the energetic crisis and their completely understanding is a crucial step for their utilization as an energy resource.

RESULTS AND DISCUSSION

Screening for Herbaspirillum rubrisubalbicans pathogenicity in Sorghum bicolor

Initially we did a screening of 58 genotypes of *Sorghum bicolor* from 3 different populations (TABLE 4-1). The screening was based on the inoculation of all genotypes with 10^8 bacteria per mL of *Herbaspirillum rubrisubalbicans* M1 on the stalk of the plants. After 7 days of inoculation was already possible to observe symptoms of disease in several leaves. The bacterium was reisolated from the symptomatic tissues 7 days after inoculation and growth in Nitrogen free NFb-Malate media. Interestingly we found 54 susceptible genotypes and just 4 resistant genotypes. The symptoms are characterized by red stripes and points of necrosis along the leaves, as reported by James 2007. The symptoms observed in the 54 genotypes are consistent between them and shown in the Figure 4-1.

TABLE 4-1 – Sorghum genotypes analyzed (ND means not definied)

Genotype	Source	Susceptibility	Genotype	Source	Susceptibility
PI655972	Kresovich	+	**Leoti**	Kresovich	+
PI562730	Kresovich	+	**PI510757**	Kresovich	+
+PI505735	Kresovich	+	**PI506069**	Kresovich	+
PI562730	Kresovich	+	**PI329311**	Kresovich	+
Leoti	Kresovich	+	**PI642998**	Kresovich	+
PI510757	Kresovich	+	**PI297130**	Kresovich	+
PI506069	Kresovich	+	**PI562730**	Kresovich	+
PI329311	Kresovich	-	**PI508366**	Kresovich	+
Grass1	Kresovich	+	**Atlas**	Kresovich	+
PI_533766	Morris	+	**PI_576434**	Morris	+
PI_655996	Morris	+	**PI_656111**	Morris	+
PI_565121	Morris	+	**PI_656015**	Morris	+
PI_597980	Morris	+	**PI_656025**	Morris	+
PI_533869	Morris	+	**PI_656023**	Morris	+
PI_540518	Morris	+	**PI_610727**	Morris	+
PI_561472	Morris	+	**PI_534133**	Morris	+
B.Tx2928	Rooney	+	**B.Tx623**	Rooney	+
B.TX3197	Rooney	+	**B.Tx623_bmr6**	Rooney	+
B.TX378	Rooney	+	**N108_B**	Rooney	+
Rio	Rooney	+	**B.Tx642**	Rooney	+
SC372	Rooney	-	**SC748-5**	Rooney	+
B.Tx644	Rooney	+	**Della**	Rooney	+
SC414_12E	Rooney	+	**SC155_14E**	Rooney	-
B.Tx645	Rooney	+	**B.AZ9504**	Rooney	+
DL1366	Rooney	+	**PI597947**	ND	+
PI656070	ND	+	**IS3220C**	ND	-
PI533831	ND	+	**PI598121**	ND	+
PI533839	ND	+	**PI609465**	ND	+
PI533878	ND	+	**PI534155**	ND	+
PI533967	ND	+	**PI656029**	ND	+
PI564163	ND	+	**PI629059**	ND	+
BR007A	EMBRAPA - BRAZIL	+			

SOURCE: The Author (2017)

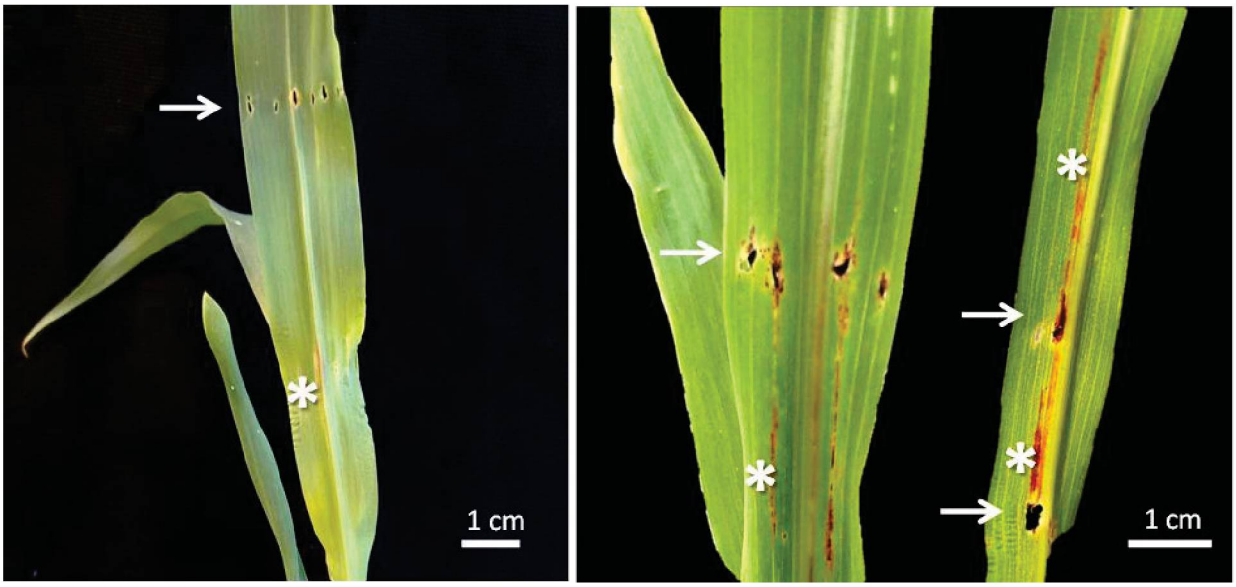


FIGURE 4-1– Pattern of symptoms observed in 54 different sorghum genotypes analyzed after the inoculation with *Herbaspirillum rubrisubalbicans* M1. The white arrows indicate the necrosis points. The white asterisk indicates the red stripes along de leave. The bacteria was inoculated on the stalk with an hypodermic needle, approximately 1 cm above the soil (not shown on the image).

Scanning electron microscopy of the epidermis of inoculated and uninoculated plants was performed to evaluate possible morphological alterations in plant tissues. For this, the plants were cultivated for 14 days, inoculated with *H. rubrisubalbicans* M1 and observed for another 7 days. At the beginning of the emergence of symptoms the leaves were collected, fixed and observed under a scanning electron microscope (FIGURE 4-2).

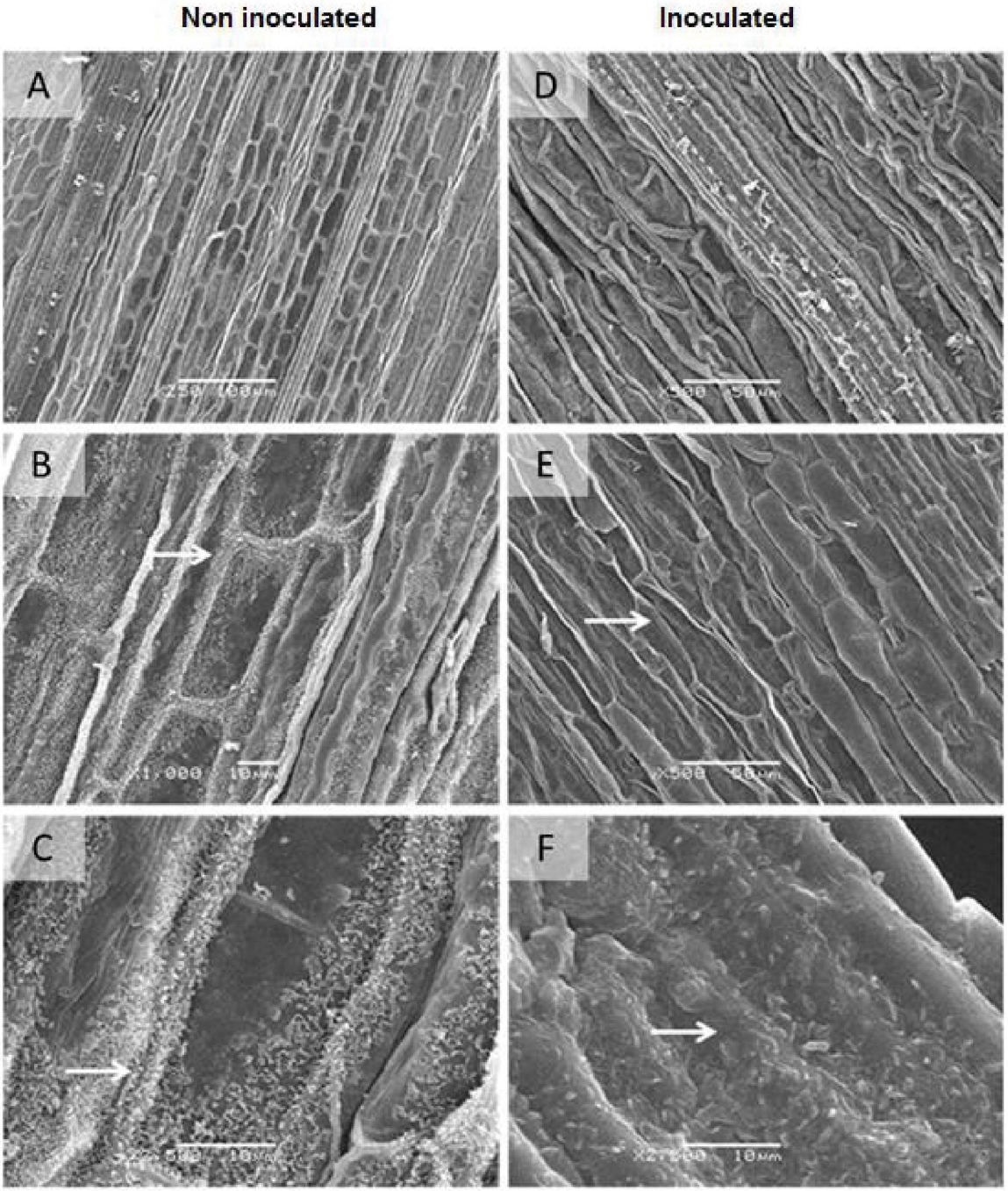


FIGURE 4-2 - Scanning Electron Microscopy of the sorghum epidermic leaves. In (A), (B) and (C) micrographs of control plants, non-inoculated with *H. rubsubalbicans* M1; (D), (E) and (F) shows micrographs of plants inoculated with *H. rubrisubalbicans* M1. White arrows indicate regions with large cuticle presence. The images were taken in a scanning electron microscope JEOL-JSM 6360 LV. The magnification used and the scales are indicated in the micrographs.

The scanning micrographs of the epidermis present morphological alterations in the intersomatic elongated cells (FIGURE 4-2). The inoculated plant cells lose turgescence and start a necrosis tissue process. In the same way, cuticular changes are observed in the

inoculated plants. The cuticle becomes thinner, less organized and smaller in comparison to the cuticle of uninoculated plants. These cuticular changes have already been described as associated to stress situations in sorghum plants (LINO et al., 2011).

To analyze the number of bacteria and how the spread of the bacteria on the tissues can be involved with the symptom development we count the bacteria in different portions of the leaf in different genotypes. All the genotypes analyzed showed the same pattern (FIGURE 4-3).

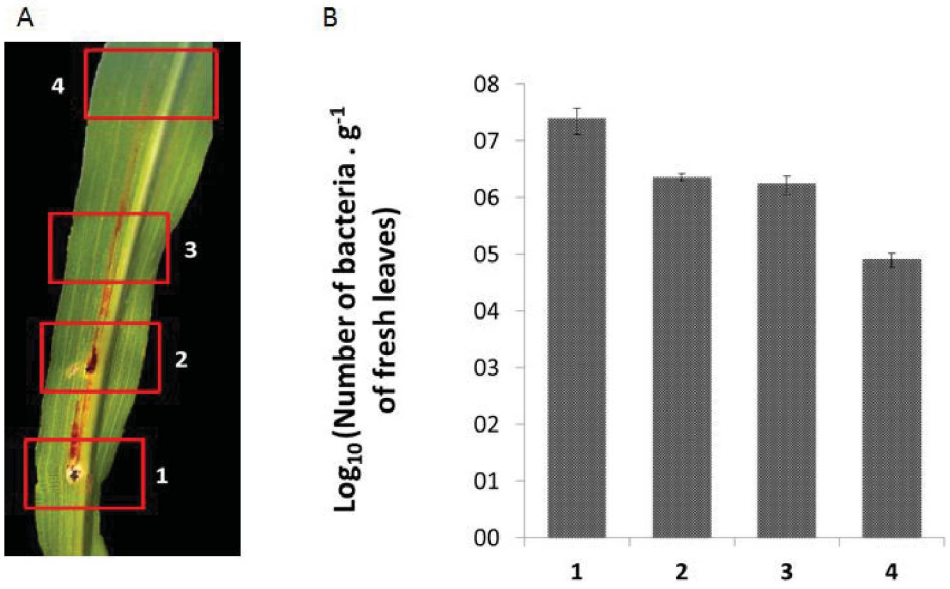
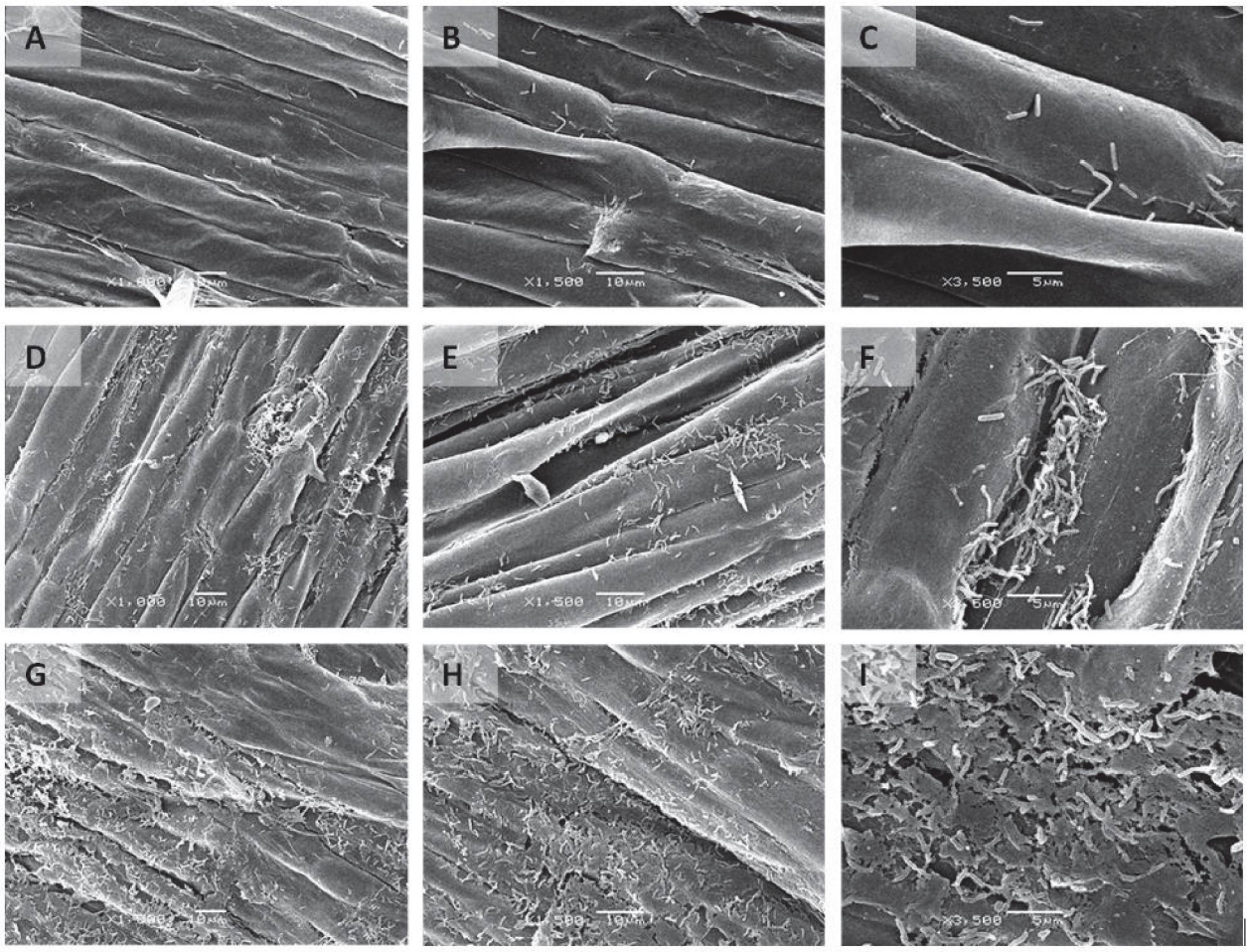


FIGURE 4-3 – Count of bacteria in different portions of the leaf. In (A) is indicated the analyzed portions of the sorghum leaf (1 to 4), where 1 is a necrotic point, 2 is an initial necrotic point, 3 is a red stripe point and 4 is an area without symptoms. In (B) the count of total bacteria inside the tissues of the specific regions.

This results support the hypothesis postulated by James (1997), where the number of bacteria and consequent blockage of the vessels is the responsible for the symptom development and necrosis. The increase in the number of bacteria and the blockage of the vessels probably affect the supply of nutrients leading to tissue damage and necrosis.

We also did an experiment doing the inoculation through the roots, and follow the association by Scanning Electron Microscopy for 1, 3 and 7 days (SUPPLEMENTARY FIGURE 4-1). Plants develop the symptoms of the disease even when the inoculation is performed by the roots.



SUPPLEMENTARY FIGURE 4-1 - Scanning Electron Microscopy of sorghum roots inoculated with *H. rubisubalbicans* M1. The sorghum plants were grown in hydroponic medium, fixed and prepared for SEM. In (A), (B) and (C) sorghum roots after 1 day of inoculation with *H. rubrisubalbicans* M1; In (D), (E) and (F) sorghum roots after 3 days of inoculation with *H. rubrisubalbicans* M1; In (G), (H) and (I) roots of sorghum after 7 days of inoculation with *H. rubrisubalbicans* M1. The images were taken in a scanning electron microscope JEOL-JSM 6360 LV. The magnification used and the scales are indicated in the micrographs.

PAMP`s involvement in the red stripe disease

Once the bacteria is in contact with the plant, they reach the extracellular surface receptors responsible for recognize conserved microbial features, the pathogen-associate molecular patterns (PAMPs). This encounter initiates the PAMP-triggered immunity (PTI) that tray to control the infection before the microorganism colonizes the plant (CHISHOLM et al., 2006). The plant response can be associated with MAP kinase signaling, transcriptional

induction of pathogen responsive genes, production of ROS and deposition of callose. All these responses help to control the infection and to protect the plant against the pathogen. However, some pathogenic organisms can alter this recognition by the secretion of effector proteins which will alter the resistance signaling or the plant resistance response (CHISHOLM et al., 2006).

Being the spread of the bacteria important for the development of the disease and based on the fact that the activation of the plant immune system is important for the protection of the plant we perform an assays pretreating the leaves with PAMP`s and evaluating the development of the disease. The plants can recognize various components of the gram-negative bacteria, including the fagellin (the protein portion of the flagellum). The central region of the flagellin is variable, but the C-terminal and N-terminal portions are highly conserved, which make this region a good PAMP. The plants can also respond to fungal molecules such as chitin and ergosterol, constitutive of the fungi cell wall. Based on that, we pretreat the leaves with flg22 (peptide with 22 highly conserved aminoacids from flagellin) and Chitin (chitooligosaccharide with 8 Carbon atoms - Chi8) to evaluate if the preactivation of the immune system response would be important for the control of the red stripe disease.

Some susceptible genotypes (Btx642, PI506069, PI508366, SCI414, BTx645, LEOTI and SCI748) were treated by infiltration with the PAMP's flg22 and Chitine (Chi8) and after 2 hours were infiltrated at the same location with 10^7 bacteria. Seven to ten days after inoculation it is possible to observe that in the regions where the pre-activation of the immune system was made there was a significant protection of plant tissues. The size of the lesion area was significantly decreased in all tested genotypes with both treatments (FIGURES 4-4 and 4-5). The measurement of the lesions area was made using K-means algorithm (SUPPLEMENTARY FIGURE 4-2).

This results shows that the pretreatment with PAMP's reduce significantly the size of the lesion promoting an efficient protection of the leaf against the pathogen. According to the experiment, the flg22 was more efficient in the most part of the genotypes when compared with the chitooligosaccharide.

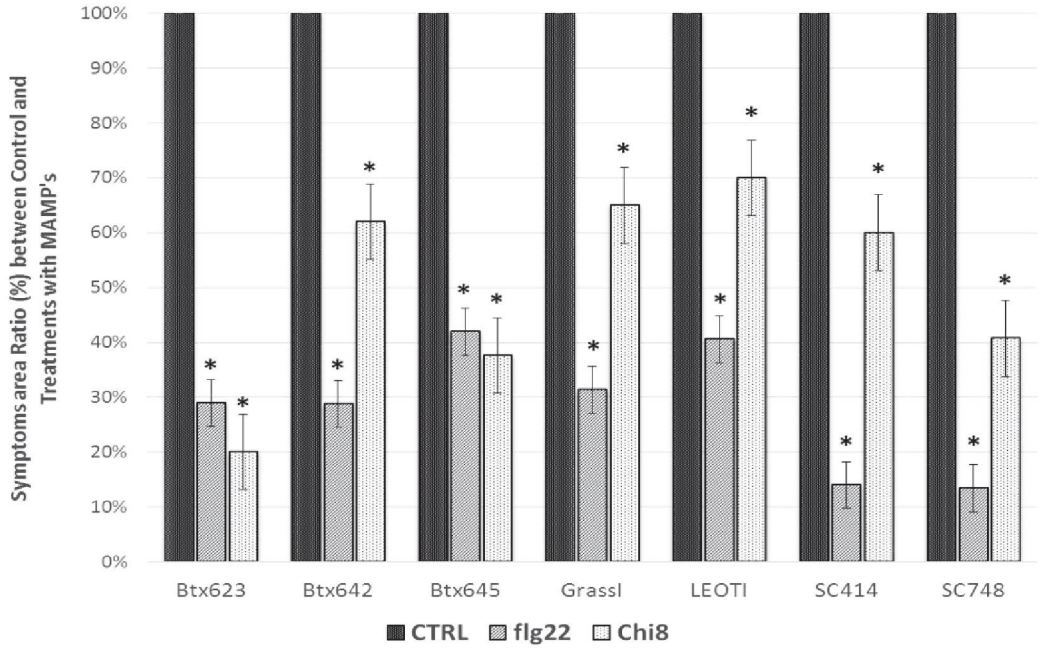


FIGURE 4-4– Ratio between the symptomatic area of control leaves (inoculated with saline solution then *H rubrisubalbicans*) and pre-treated MAMP`s leaves (inoculated with flg22/Chitin then *H. rubrisubalbicans*).

.

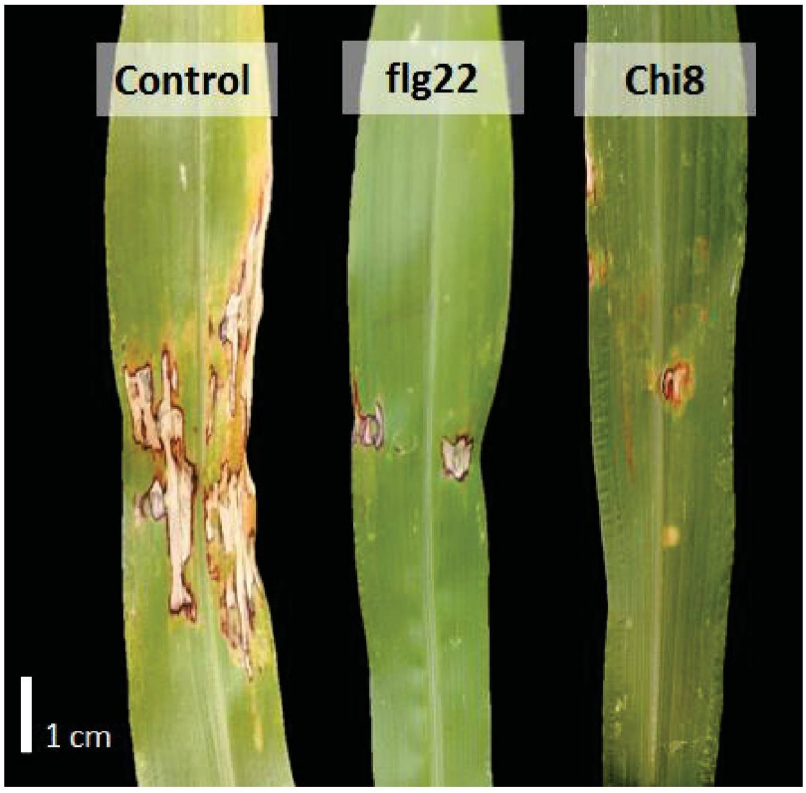
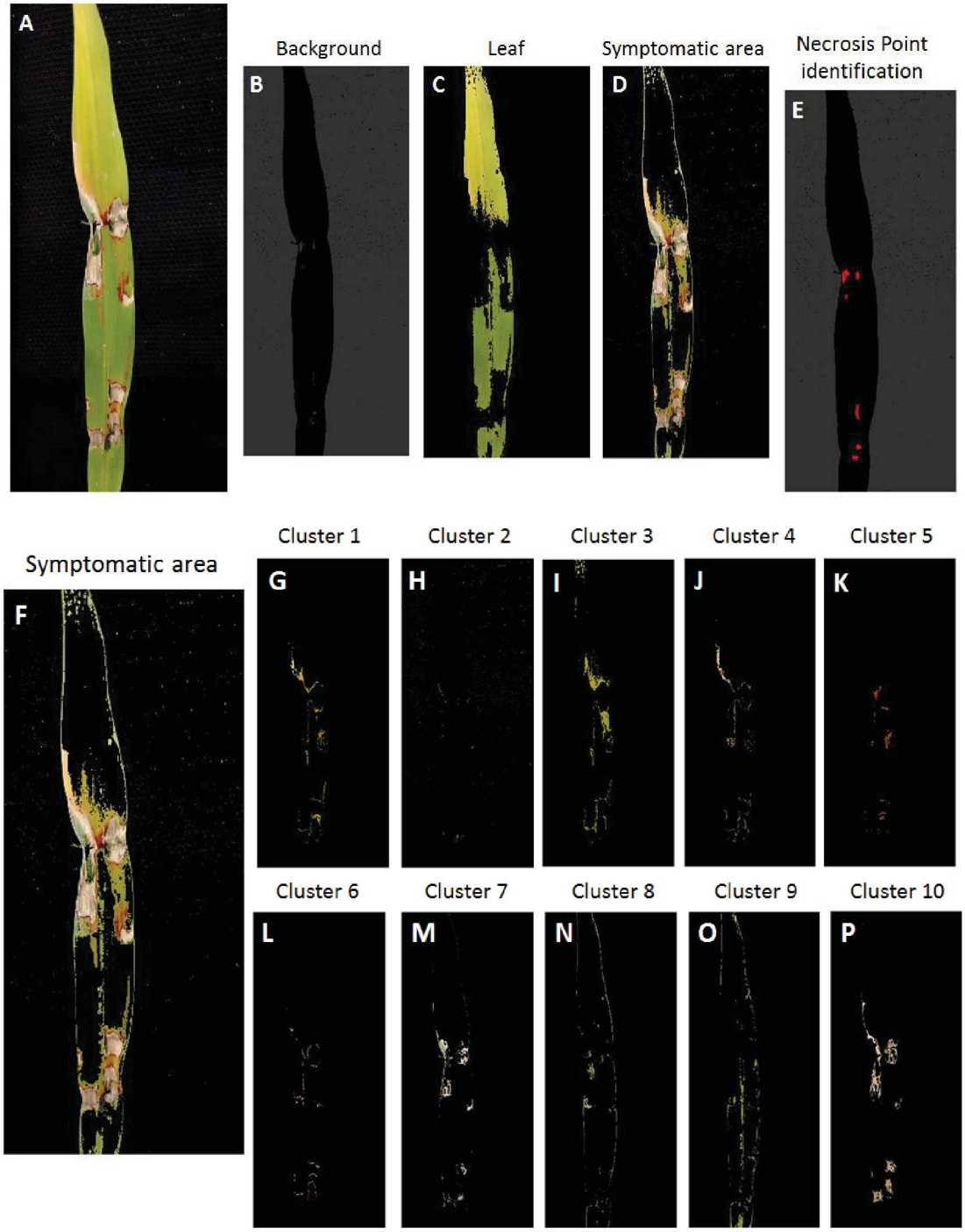


FIGURE 4-5 - Sorghum leaf (Btx632) inoculated with mock solution and *H. rubrisubalbicans* M1 (Control), with 100 nM of flg22 and *H. rubrisubalbicans* M1 (flg22) and with 1 µM of Chitin and *H. rubrisubalbicans* M1 (Chi8).



SUPPLEMENTARY FIGURE 4-2 - Measurement of the lesions on sorghum leaves. The measurement was made using the K-means algorithm. In (A) the original picture analyzed. In (B), (C) and (D), the original picture was divided in 3 different portions: Background (B), healthy leaf (C) and symptomatic area (D). Then, the necrosis identification (E) was made based on the background. The points of necrosis were identified and counted as symptomatic area. In (F), the symptomatic area was showed again and was divided in different clusters (G to P), representing the different colors patterns. Some clusters, as 8 and 9 (O and P) are just the edges of the leaves and are not real symptoms, so this clusters, as well as Cluster 3 (I) were removed from the analysis. This approach allowed a more refined analysis of the symptomatic and healthy areas.

To analyze if the preinfiltration with the PAMP`s is capable of modulate de bacteria colonization inside the leaf tissue we perform a count of bacteria in different portions of the leaves from different genotypes (FIGURE 4-6 and SUPPLEMENTAR FIGURE 4-3), starting at the inoculation point and following the spread of bacteria along the leaf. The result shows that even flg22 and chitooligossacharide are capable of modulate the bacterial colonization. The number of bacteria decrease significantly as the count gets further away from the inoculation point.

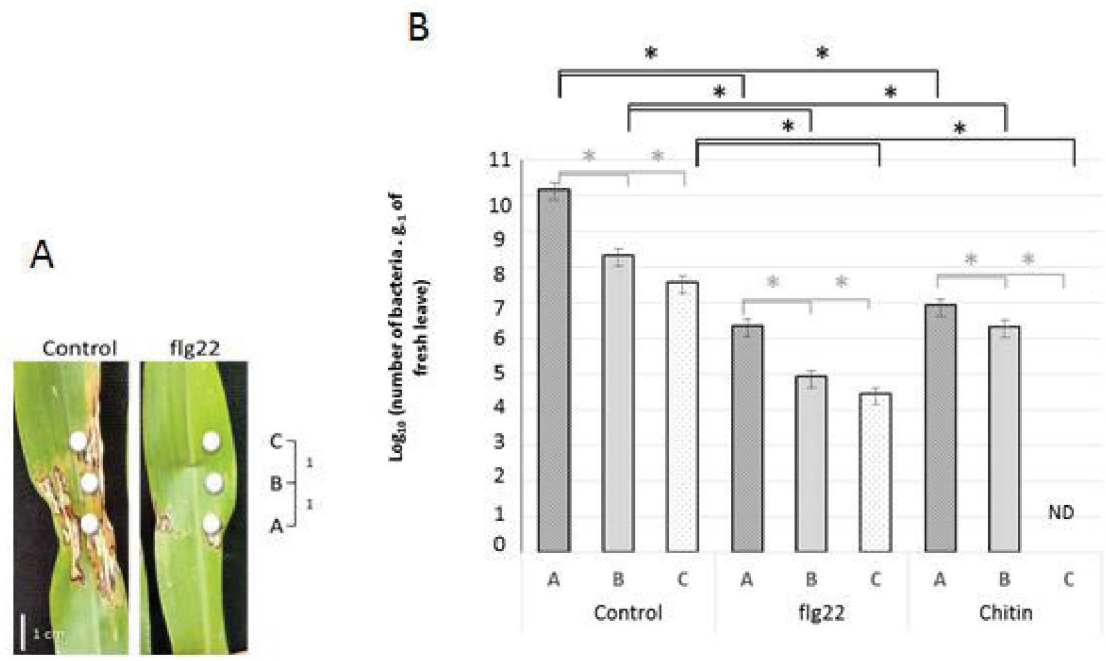
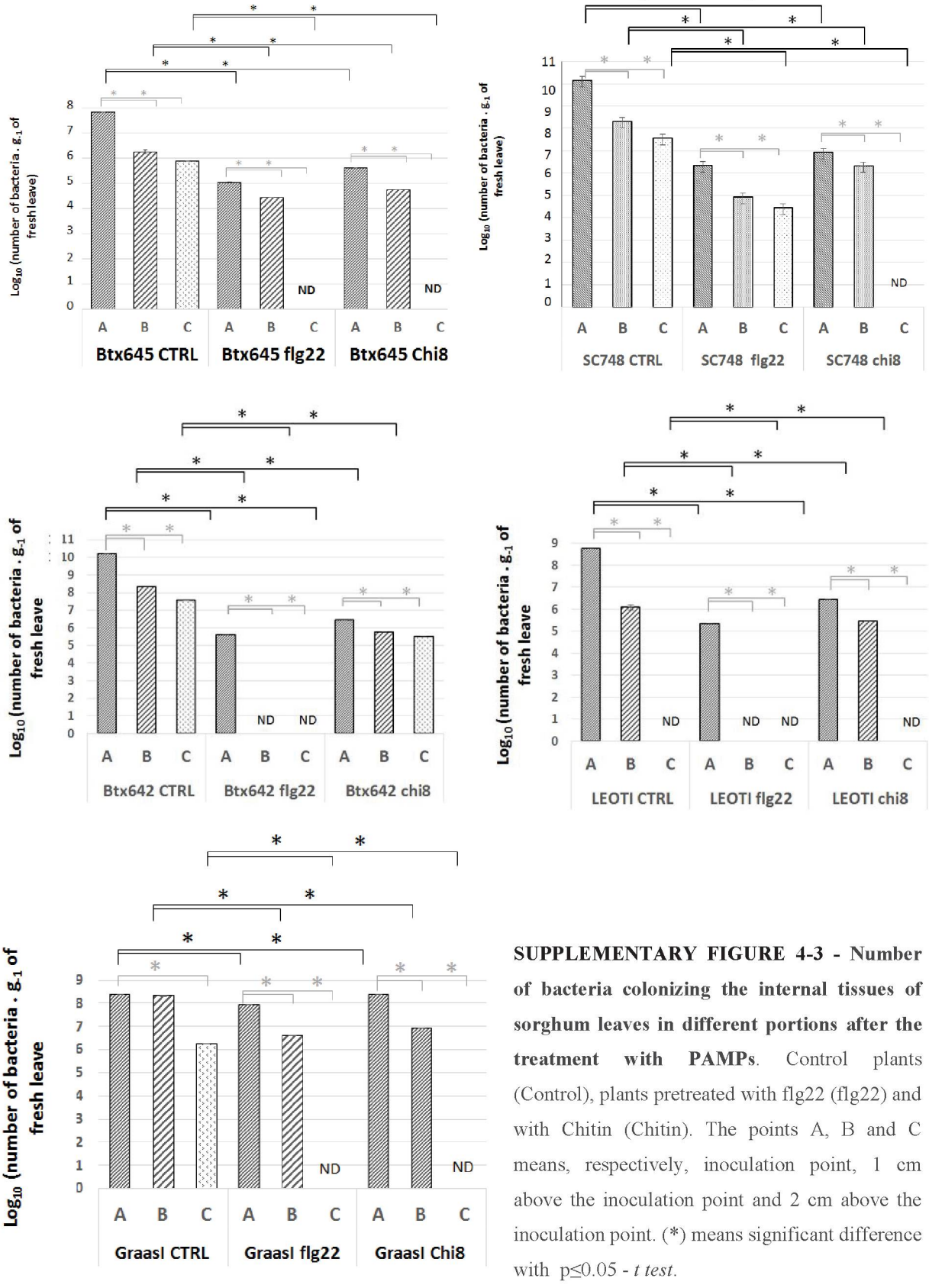


FIGURE 4-6 – Count of bacteria in different leaf portions after the treatment with PAMPs. (A) Schematic demonstration of analyzed points. (B) Graph showing the number of bacteria colonizing the internal tissues of sorghum leaves (Btx623) in different portions around the inoculation point from the Control plants (Control), plants pretreated with flg22 (flg22) and with Chitin (Chitin). The points A, B and C means, respectively, inoculation point, 1 cm above the inoculation point and 2 cm above the inoculation point. (*) means significant difference with $p \leq 0.05$ - *t test*.



SUPPLEMENTARY FIGURE 4-3 - Number of bacteria colonizing the internal tissues of sorghum leaves in different portions after the treatment with PAMPs. Control plants (Control), plants pretreated with flg22 (flg22) and with Chitin (Chitin). The points A, B and C means, respectively, inoculation point, 1 cm above the inoculation point and 2 cm above the inoculation point. (*) means significant difference with $p \leq 0.05$ - *t test*.

According to this experiment, the number of bacteria in the inoculation point is significantly reduced when the leaf is pretreated with flg22 or chitooligosaccharide. The same occurs with the spread of bacteria around the leaf. The colonization in the points B and C is also reduced compared with the controls. This means that the PAMP`s may have a function in control the internal leaf bacterial colonization.

The recognition of pathogen is the first step to plant starts the response through the PAMPs-triggered immunity. One of the associated responses is the callose deposition. This deposition is important to reinforce the cell wall close to the sites of infection and control the growth of the microbial (CHISHOLM et al., 2006).

To understand how the pretreatment with PAMP`s can have function in control the number of bacteria we analyzed the callose deposition. Since the callose is a physic barrier against the pathogen and the inoculation with the PAMP`s can induce this deposition by itself, seems to be interesting analyze wheter the preinitiation of the callose deposition protect the plant and avoide bacterial spreading. To understand this, first we analyze the callose deposition in control leaves (leaves inoculated just with the bacteria and presenting the red stripe symptoms). We inoculated 7 days old plants by infiltration with *Herbaspirillum rubrisubalbicans* and after 5 to 7 days later we analyze the callose deposition on the leaves (FIGURE 4-7).

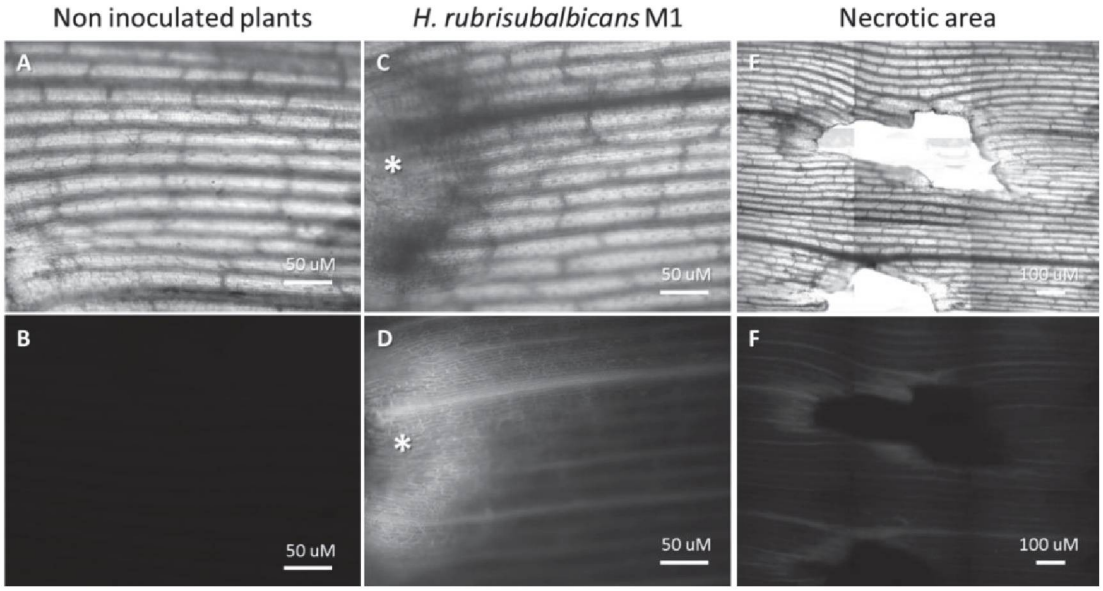


FIGURE 4-7 – Callose deposition in inoculated plants with *Herbaspirillum rubrisubalbicans* M1. (A) and (B) shows no inoculated plants, (C) and (D) shows 14 days old plants inoculated with *Herbaspirillum rubrisubalbicans* M1 and presenting symptoms of red stripe disease (*). (E) and (F) shows a scan on the necrotic area caused by *H. rubrisubalbicans*. (A, C and E) are Brightfield images; (B, D and F) are DAPI filter images.

Through this experiment it is possible to see that there is callose accumulation in the plant tissues when the bacteria are colonizing the leaves. The callose is concentrated in the symptoms points and around the necrotic sites (FIGURE 4-7E and 4-F).

To understand if the early callose deposition is the mechanism behind the PAMP`s bacteria control we perform the same experiment with plants treated with flg22 and chitooligosaccharide. We used 14 days old plants and we infiltrate this plants with flg22, chitooligosaccharide and *Herbaspirillum rubrisubalbicans*, individually. After 24 hours we stain this leaves for callose detection (FIGURE 4-8).

We can observe on this results that there is a large callose deposition between the cells in the flg22/chitooligosaccharide treatments and this deposition is more controlled in the plants inoculated just with the bacteria. When there is just bacteria inoculation we can see some points of callose deposition, but this points are concentrated just in a few cells. This can be one of the factors responsible for the smaller number of bacteria inside the tissues when we pretreat the plants with the PAMP`s and then challenge with the bacteria. The earlier callose deposition can contribute to the control of the bacterial colonization.

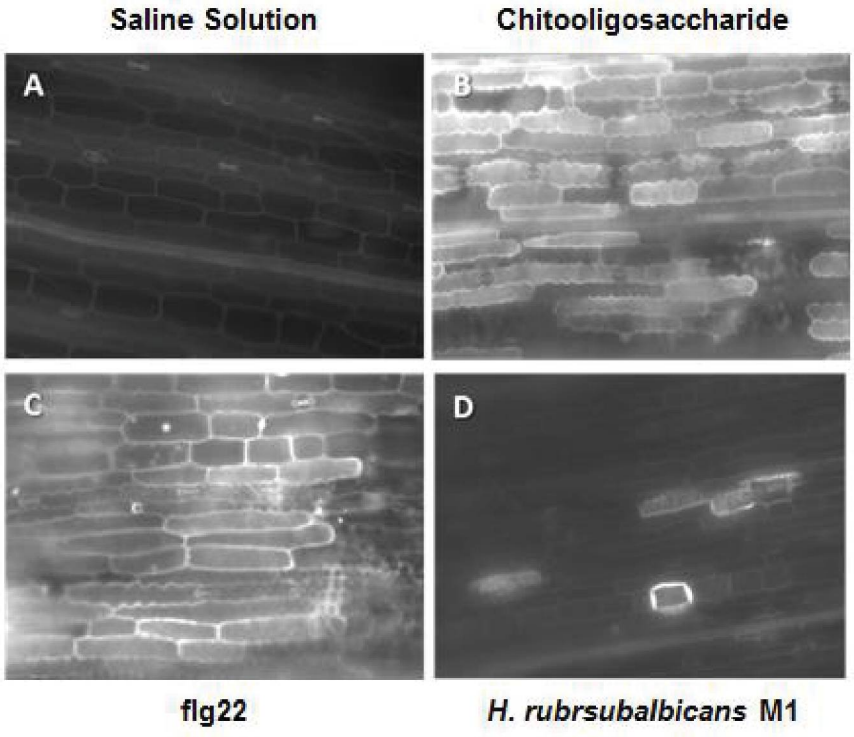


FIGURE 4-8 – Callose deposition in 14 days old plants. The plants were infiltrated with saline solution (A), chitooligosaccharide (B), flg22 (C) and *Herbaspirillum rubrisubalbicans* M1 (D) and were observed in fluorescence microscope 24 hours after infiltration. The fluorescence indicates the callose accumulation between the cells.

ABA pre infiltration

Olivares (1997) and James (1997) reported that the substomatal cavities of sugarcane were colonized during the infection with *Herbaspirillum rubrisubalbicans* and this can suggest that these points can be an entry or an invasion site for the bacteria colonization. Also this kind of colonization could allow the scape of the bacteria to infect neighborhood plants or neighborhood sites in the same plant. In contrast, James (1997), found that in Sorghum the colonization is more localized in the vascular system and there is no adjacent bacterial cells in other tissues.

To understand if the bacteria can use the sorghum stomata structure to enhance the colonization we perform an experiment using abcisic acid (ABA) to close the stomata during the infection (FIGURE 4-9).

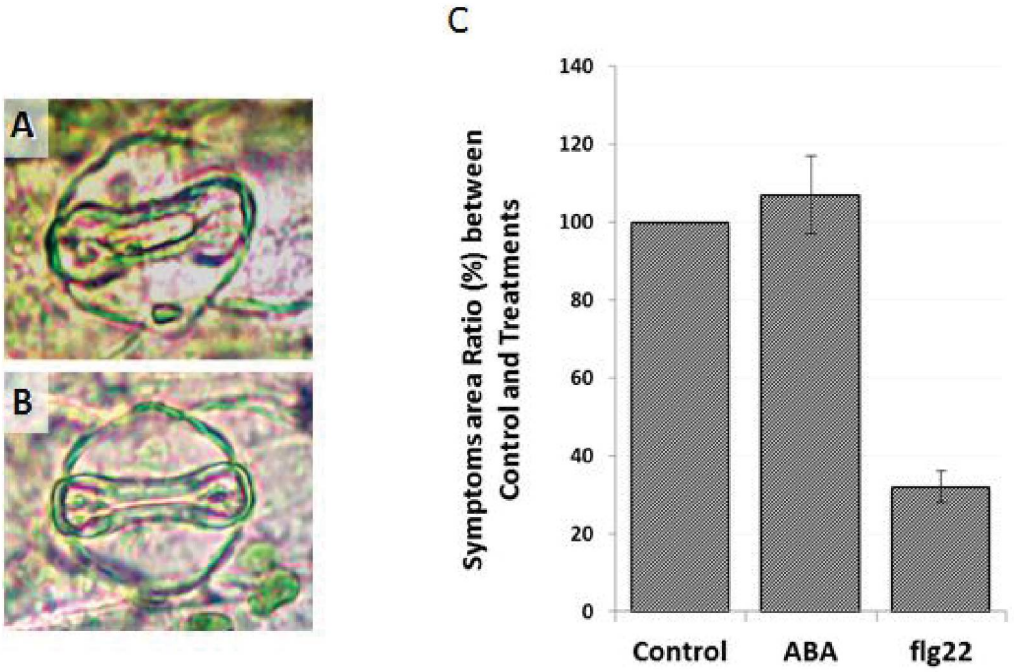


FIGURE 4-9 – Avaluation of stomata involvement in the red stripe disease development. Seven days old sorgum plants were pretreated with abcisic acid (ABA) and then were inoculated by infiltration with *Herbaspirillum rubrisubalbicans*. In (A) the stomata open, In (B) the stomata closed after the ABA infiltration, in (C) the symptome area Ratio (in porcentage) between the control plants, the plants pre treated with ABA and plants pre treated with flg22.

This results shows that the stomata are not crucial to the infection and the main bacterial colonization may occur by the internal tissues, as the vassels, since there was no

difference in the development of the symptoms or in the count of bacteria when the stomata was closed.

QTL analysis

Knowing that the immune system is a crucial component to the development of the disease and that we found 4 genotypes resitent to the red stripe disease we looked for another important factors that could be involved with this interaction. We analyzed a RIL population from a crossing between a susceptible genotype (BTx623) and a resistent genotype (SC155-14E). In this analysis we focused in the susceptiblity to the red stripe disease using a patogenicity scale from 0 to 5, where 0 means no symptome and 5 means high level of symptome (SUPLEMENTARY FIGURE 4-4) and counts of total bacteria colonizing the leaf tissues.

SUPPLEMENTARY FIGURE 4-4 – Sorghum red stripe disease arbitrary patogenicity scale. The scale goes from 0 to 5, where 0 means no symptome and 5 means high level of symptomes

Disease Testing.

A significant variation between the RILS was observed in both analyses (SUPLEMENTARY TABLE 4-1). Bacteria counts varied from 0 to 10^8 bacteria per gram of fresh leaf, while pathogenicity scores varied from no symptom to 4 (FIGURE 4-10). Significant differences were identified for all experimental factors, including RIL, experimental run, replication, and the interaction of RIL with replication and experimental run

for both visual disease response and bacterial counts (TABLE 2). Pearson correlation coefficients between replications within runs for visual ratings were high (P=0.72-0.88) and the correlation between runs was 0.64 (*data not shown*). For bacterial counts, correlation between replications within runs ranged from 0.82-0.90 and between runs was 0.80 (*data not shown*). A strong correlation (P=0.74) between both analyses was observed, indicating that the high number of bacteria is correlated with the high levels of pathogenicity (FIGURE 4-11)

SUPPLEMENTARY TABLE 4-1 - Adjusted means for parents, BTx623 and SC155-14E, and RILS average, minimum, and maximum trait values for visual disease response and bacterial counts calculated within and across experimental runs.

	Visual Disease Response				Bacterial Counts			
Experimental Run	1	2	3	All	1	2	3	All
BTx623	4.00	3.00	3.67	3.56	7.21	7.24	6.57	7.01
SC155-14E	0.00	0.00	0.00	0.00	2.75	3.27	2.38	2.80
RIL Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RIL Maximum	5.00	4.00	4.67	4.11	8.66	7.82	7.17	7.58
RIL Average	2.16	1.98	1.81	1.98	4.70	4.92	4.40	4.67

SOURCE: The Author (2017)

TABLE 4-2 - Analysis of variance significance values for 103 RILs evaluated in a greenhouse setting for their *H. rubrisubalbicans* visual disease response and bacterial count.

H. rubrisubalbicans Visual Disease Response					
Source	DF	Type III SS	Mean Square	F Value	Pr > F
RIL	105	953.14	9.08	31.28	<.0001
Replication	2	31.66	15.83	54.55	<.0001
Run	2	19.03	9.52	32.80	<.0001
RIL*Replication	210	76.57	0.36	1.26	0.0259
RIL*Run	210	207.19	0.99	3.40	<.0001
Replication*Run	4	1.24	0.31	1.07	0.3703
H. rubrisubalbicans Bacterial Counts					
Source	DF	Type III SS	Mean Square	F Value	Pr > F
RIL	105	5444.86	51.86	67.73	<.0001
Replication	2	38.63	19.32	25.23	<.0001
Run	2	43.68	21.84	28.52	<.0001
RIL*Replication	210	205.27	0.98	1.28	0.0187
RIL*Run	210	553.47	2.64	3.44	<.0001
Replication*Run	4	4.72	1.18	1.54	0.1896

SOURCE: The Author (2017)

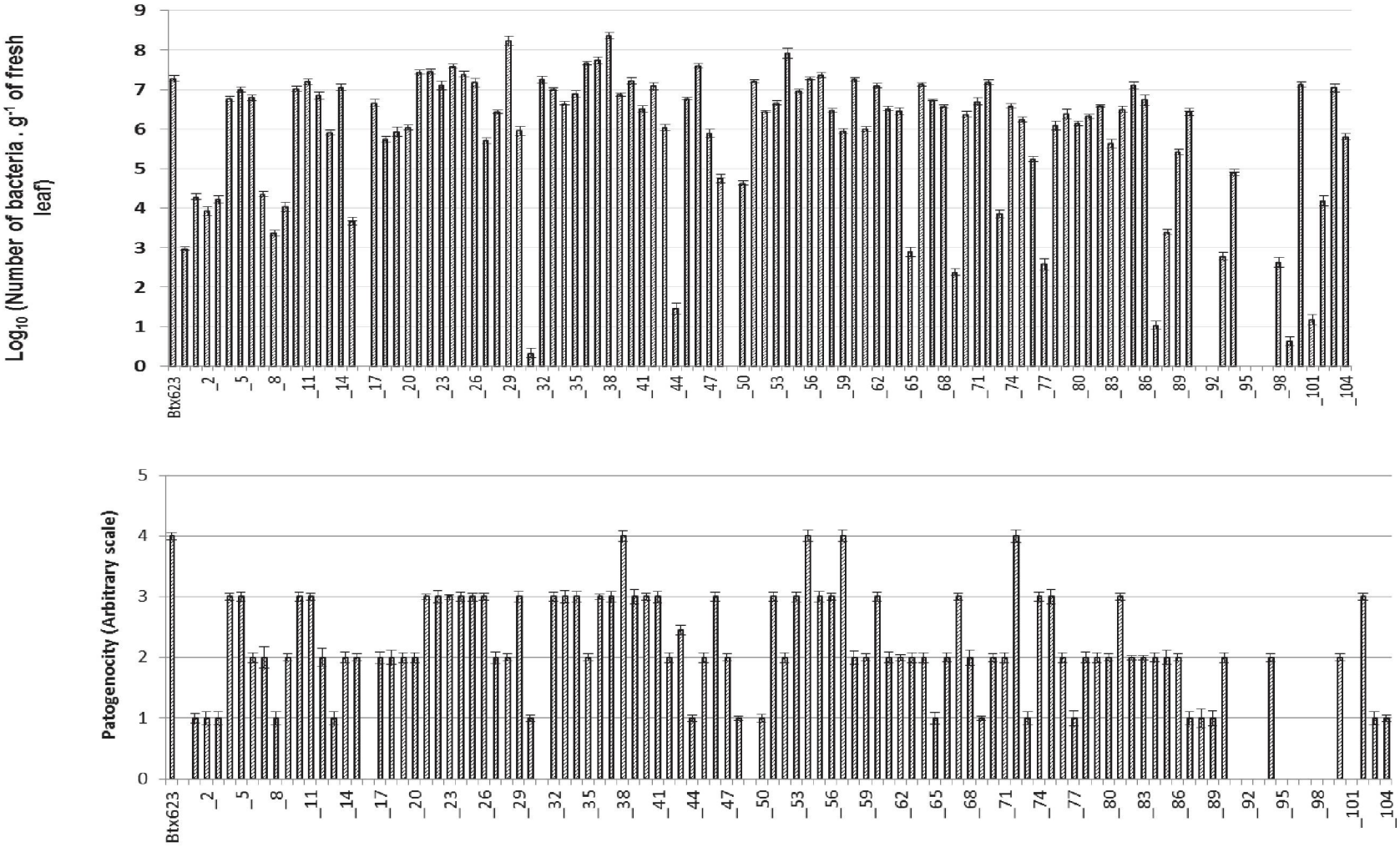


FIGURE 4-10 – Analysis of 102 RILs from the crossing between BTx623 (susceptible) and SC155-14E (resistant). In (A) the count of bacteria after 7 days of inoculation. In (B) the pathogenicity scale of the population. The bars indicate the standard error.

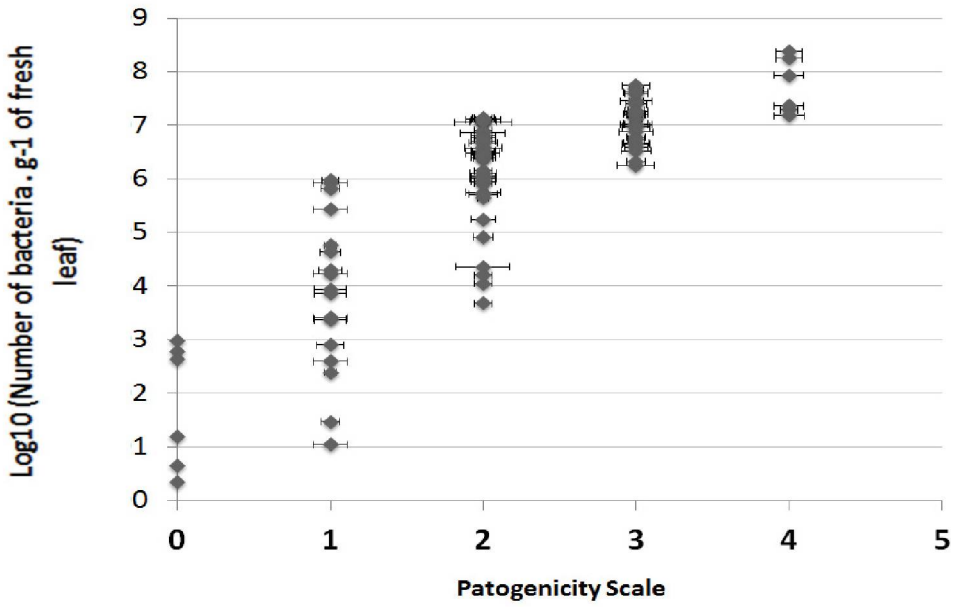


FIGURE 4-11 – Correlation between the Number of bacteria inside the plant tissues and the pathogenicity observed in the RIL population (Btx623 x SC155-14E).

QTL Identification.

WinQTL cartographer was used to identify putative QTL for *H. rubrisubalbicans* response in a BTx623 x SC155-14E RIL mapping population. QTL for the visual ratings were detected on chromosomes 3 and 7 within and across experimental runs with the exceptions that run 1 did not detect QTL on chromosome 7 and run 2 did not detect QTL on chromosome 3 (TABLE 4-3, *only QTL across runs shown*).

QTL for bacterial counts were detected on chromosomes 7 and 10 within and across experimental runs (TABLE 4-3, *only QTL across runs shown*). Two QTL located next to one another on chromosome 7, which should most likely be considered one QTL, explained 15.11-17.59% of the variation in bacterial counts, and overlapped with QTL detected using visual ratings (FIGURE 4-12). The BTx623 allele for this QTL lowered bacterial counts by 0.86-0.96 on a scale of 0-8.66 (scale based on bacterial counts after log transformation). QTL located on chromosome 10 explained 20% of the variation in bacterial counts. The SC155-14E allele for this QTL lowered bacterial counts by 1.18.

The analysis indicates a region on the chromosome 7 as an important region for resistance/pathogenicity. Analysis of this region showed an important group of genes that are closely related with the plant response against pathogens, such as CERK1, (a chitin receptor involved with the PAMP-triggered immunity), RIN4 and PtiI involved with the response

related do ETI (Effector Triggered Immunity), CML and CDPK (both related to the calcium signaling pathway).

TABLE 4-3 - Characteristics of QTL detected using WinQTL cartographer for *H. rubrisubalbicans* visual disease response and bacterial counts across three experimental runs in a BTx623 x SC155-14E RIL mapping population.

QTL #	Parameter	Chr	Peak Position (cM)	Peak Position (bp)	LOD	Additive effect	R2	Support Interval (cM) 2 LOD	Support Interval (bp) 2 LOD
1	Visual Disease Response	3	44.81	8,313,065	2.65	-0.31	0.1015	41.9-50.0	7,660,601-9,976,473
2	Visual Disease Response	7	69.41	58177830	2.67	0.32	0.1040	59.1-79.7	52,889,017-89,787,318
1	Bacteria Count	7	69.41	58,177,830	5.11	0.96	0.1759	65.7-73.9	57,112,406-58,841,459
2	Bacteria Count	7	77.61	59,470,024	4.49	0.86	0.1511	73.9-79.6	58,841,459-59,787,247
3	Bacteria Count	10	36.21	4,234,323	5.57	-1.18	0.2007	27.1-38.5	2,897,335-4,740,139

SOURCE: The Author (2017)

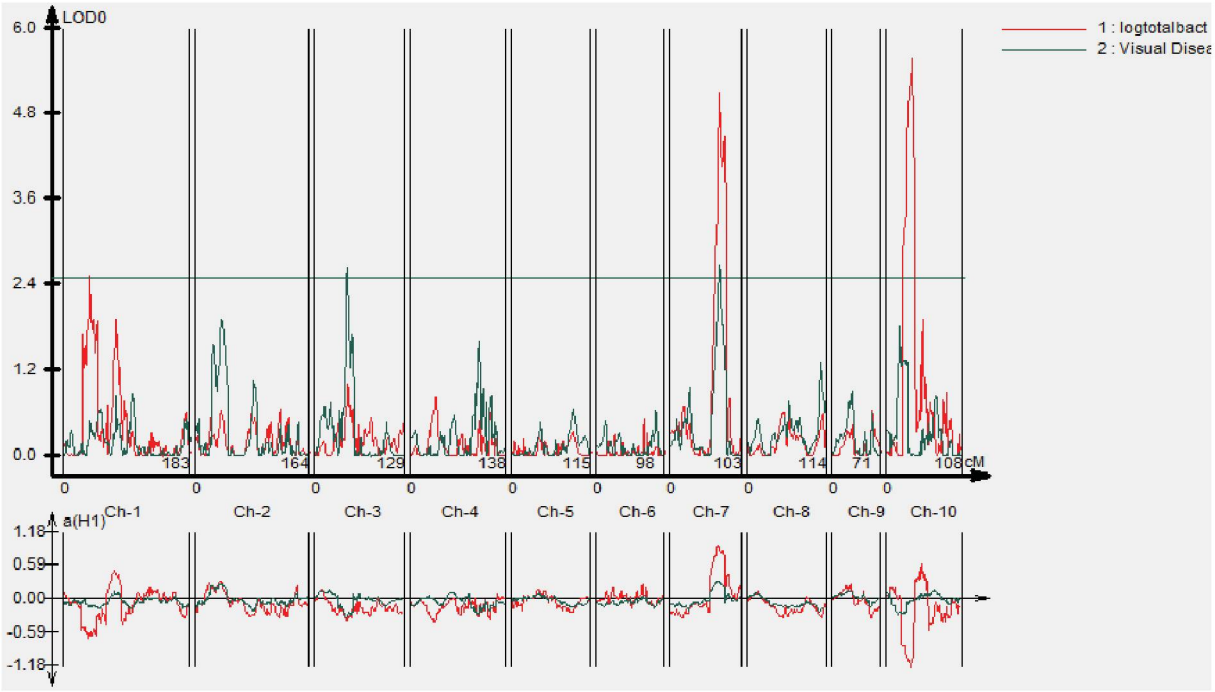


FIGURE 4-12 - Curves of LOD values (upper graph) and additive effects (lower graph) of QTL detected using visual disease response and bacterial counts in a BTx623 x SC155-14E RIL mapping population.

Being this population already mapped, we were able to do a QTL analysis of this data. The count of number of bacteria had a variation between the samples, so this QTL analysis was not satisfactory. Being this experiment an open assay in greenhouse this variation was expected. In contrast, the assay using the pathogenicity scale data was consistent and was statistically acceptable with a good correlation between the replicates.

The QTL analysis indicate a region (Position 70.81) on the Chromosome 7 as an important region for the pathogenicity. Analysis of this regions shows an important group of genes that are closely related to the plant response against pathogens, such as CERK1, a chitin receptor involved with the PAMP-triggered immunity, RIN4 and PtiI involved with the response related do ETI (Effector Triggered Immunity), CML and CDPK, proteins related to the Calcium signaling pathway.

Transcriptomic Analysis

We perform then an RNAseq analysis in Sorghum leaves inoculated with *Herbaspirillum rubrisubalbicans* M1 and presenting the symptoms of disease to identify potential factors involved with the pathogenicity interaction. We growth the plants in an inert soil for 7 days then we inoculated this plants with a hypodermic needle in the stalk with 10^7 bacteria. Seven days after inoculation the plants presented symptoms of the red stripe disease and the symptomatic leaves were harvested to the RNA extraction. The experiment was made in triplicates and we compare non-inoculated plants (P) with Inoculated Plants (PB – Plant+bacteria). The Supplementary Table 4-2 shows the yiel of the RNA purification.

SUPPLEMENTARY TABLE 4-2 – Yield of the Sorghum leaves RNA extraction

Samples	Control Plants			Inoculated Plant		
	P1	P2	P3	PB1	PB2	PB3
Initial amount of tissues (mg)	212,9	250,3	217,2	193,9	264,3	154
Total RNA yield (µg)	112	36	61	9,7	58,9	18,1
RNA yield after LiCl precipitation (µg)	19,9	12,2	10,5	9,1	20,7	11,9

SOURCE: The Author (2017)

The extracted rRNA was depleted and the resulting mRNA was used for the construction of the cDNA libraries. Transcriptomic libraries were constructed and sequenced on the Ion Proton ™ System - Life Technologies Next Generation Platform. After the sequencing, the trimming and mapping of the samples was performed in the CLC Genomics Workbench 7 software using the Sorghum bicolor genome (PETERSON, et al. 2009). The Table 4-4 shows the number of reads of each replicate and the size average of the reads. The Table 4-5 shows the percentage of mapping with ribossomic RNA, mithocondria, chloroplast and coding genome of Sorghum.

TABLE 4-4 – Number of reads in each library and average of the reads size.

	Total number of reads	**Number of reads after Trimming**	**Average read size**
P1	21589003 (100%)	21266660 (98,50%)	77,16
P2	10457112 (100%)	10329559 (98,78%)	82,26
P3	19035661 (100%)	18767322 (98,59%)	82,3
PB1	23649079 (100%)	23326723 (98,63%)	88,26
PB2	17718281 (100%)	17563323 (99,12%)	86,08
PB3	16193114 (100%)	16012630 (98,88%)	86,36

SOURCE: The Author (2017)

TABLE 4-5 –Mapping of the sequenced libraries against the *Sorghum bicolor* genome

	P1	**P2**	**P3**
Total of reads after trimming	21266660 (100%)	10329559,00 (100%)	18767322 (100%)
rRNA	6373877 (29,97%)	3015071 (29,19%)	4775149 (25,44%)
Mithocondria	73068 (0,34%)	31507 (0,31%)	69019 (0,37%)
Chloroplast	1116362 (5,25%)	492985 (4,77%)	1007489 (5,37%)
Sorghum coding Genome	10648598 (50,07%)	5220611 (50,54%)	10282440 (54,79%)
	PB1	**PB2**	**PB3**
Total de reads after trimming	23326723 (100%)	17563323 (100%)	16012630 (100%)
rRNA	5054653 (21,66%)	3000343 (17,08%)	2806366 (17,52%)
Mitocondria	102830 (0,44%)	76968 (0,44%)	70732 (0,44%)
Cloroplasto	834999 (3,58%)	692386 (3,94%)	582687 (3,64%)
Genoma Sorghum	13302357 (57,02%)	10611434 (60,41%)	9874134 (61,66%)

SOURCE: The Author (2017)

Cluster analysis of the samples (FIGURE 4-13A) and principal component analysis (PCA) were also performed (FIGURE 4-13B) to verify the reliability of the data obtained. According to the data, the replicates are grouped and the treatments are distanced, showing a good resolution and correlation of the obtained data.

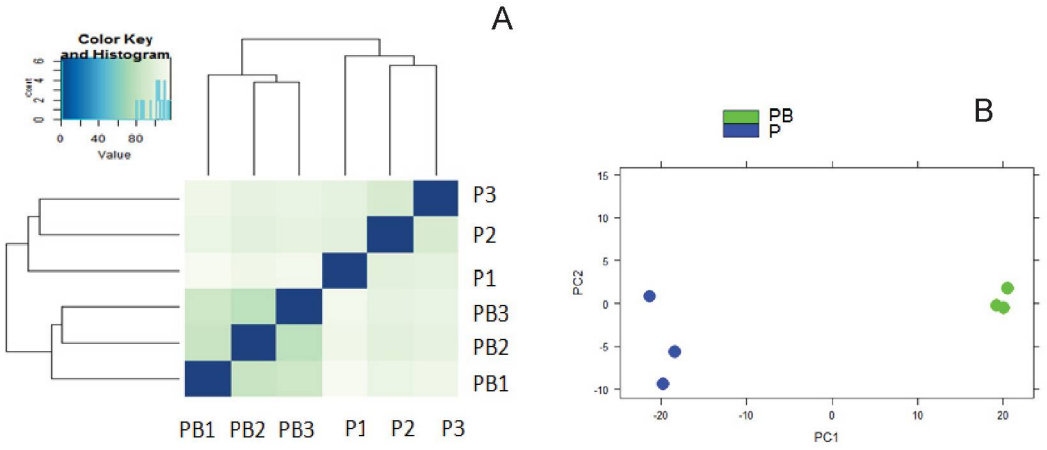
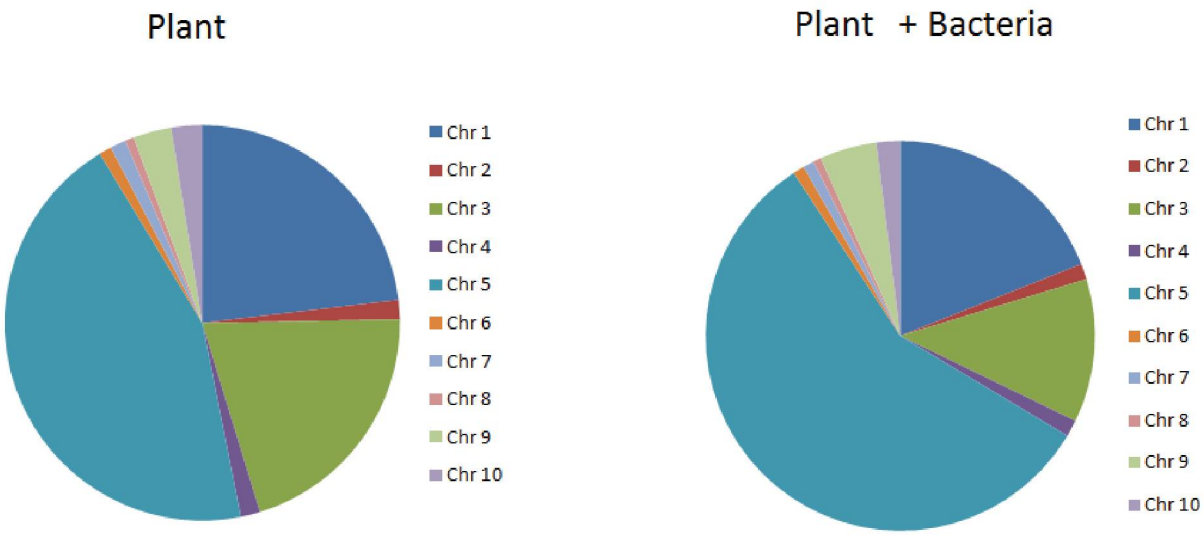


FIGURE 4-13 – (A) Cluster analysis between samples and in (B) Principal component analysis (PCA) between transcriptome samples of sorghum leaves inoculated (PB) and non inoculated (P) with *H. rubrisubalbicans*.

After mapping, the distribution of the reads mapped with the sorghum genome was analyzed according to the location in the different chromosomes (SUPPLEMENTARY FIGURE 4-5). The data shows that the distribution of the reads in the chromosomes was different among the treatments. Chr1 had more genes mapped in the non inoculated library, otherwise the Chr 3 had fewer expressed genes during interaction with *H. rubrisubalbicans*. These data suggest that the presence of the bacteria colonizing the internal tissues causes an alteration of the global gene expression in sorghum.



SUPPLEMENTARY FIGURE 4-5 - Analysis of the distribution of the reads mapped with the sorghum genome according to the location in the different chromosomes.

The libraries had a mean coverage of 2.8 for the uninoculated leaves and 4.2 for the inoculated leaves. For all the further analyzes the genes with coverage equal to or greater than 1 (approximately 6,000 genes) were considered.

Differentially expressed genes were then analyzed between treatments using the DESeq and edgeR packages of R Program. Only the genes found in the two analyzes with $p \leq 0.05$ were considered as differentially expressed. We identify 380 differentially expressed genes, being 193 Up-regulated and 187 down-regulated (SUPPLEMENTARY TABLE 4-3). Among the overexpressed genes in the inoculated plants, the great majority corresponds to hypothetical proteins, as shown in Figure 4-14.

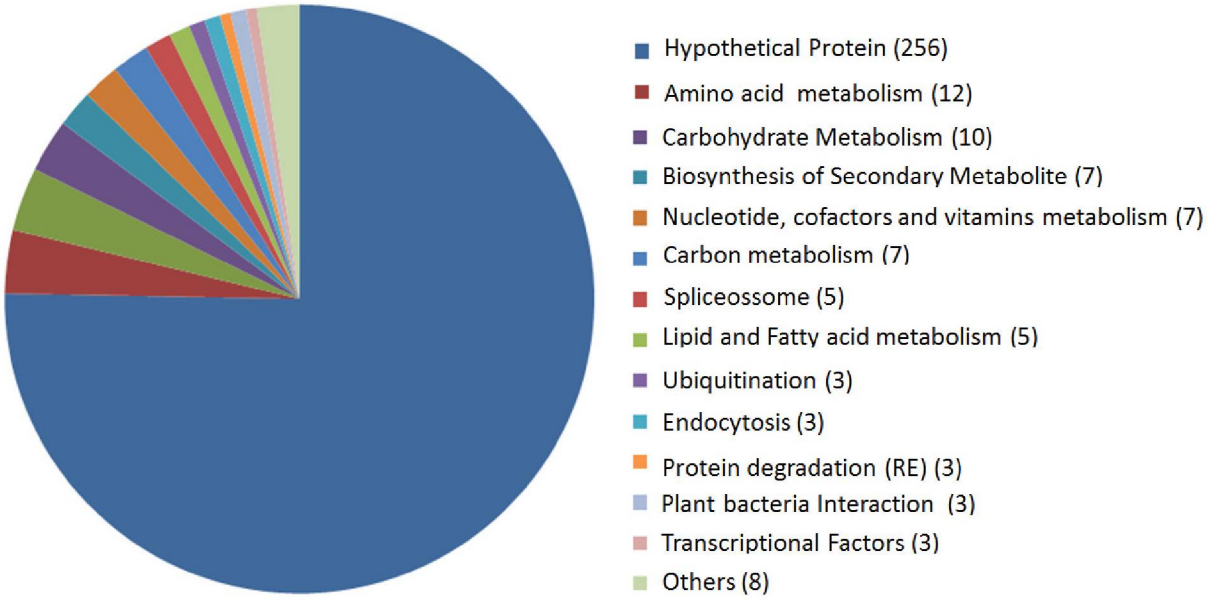


FIGURE 4-14 - Metabolic category distribution of Up-regulated genes during the analysis of the transcriptomic library of sorghum leaves inoculated with *H. rubrisubalbicans* M1.

One of the Up-regulated genes was the gene *SORBIDRAFT_02g002150,* with a fold change of 6, this gene codes for a hypothetical protein probably related to response against pathogens. Keeg analysis shows that this protein has homology with the PR1 protein and probably has antimicrobial properties, responsible for impair the growth and propagation of pathogens. This pathway is related to the response activated by bacterial flagellin flg22. These data support the hypothesis that activation of the plant immune system is present during the interaction between *H. rubrisubalbicans* M1 and sorghum.

Among repressed genes, the most part also code for hypothetical proteins (127 genes), followed by proteins related to carbohydrate metabolism (11 genes) (FIGURE 4-15).

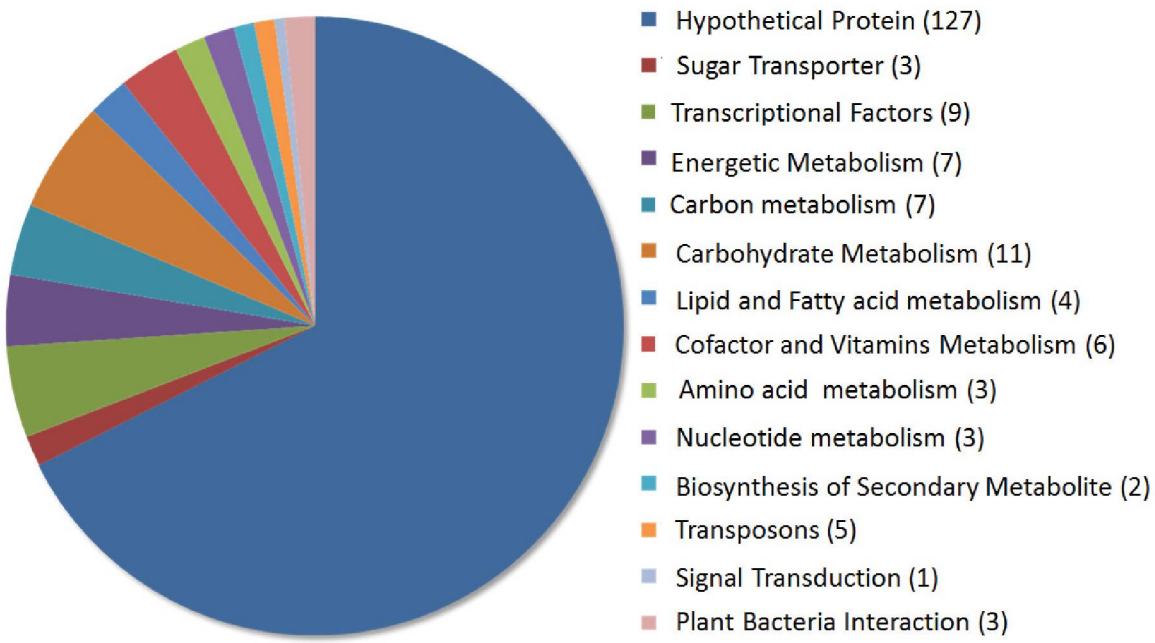


FIGURE 4-15 - Metabolic category distribution of Down-regulated genes during the analysis of the transcriptomic library of sorghum leaves inoculated with *H. rubrisubalbicans* M1.

The gene *SORBIDRAFT_10g030240*, with fold change of 2, is part of a plant-pathogen interaction pathway that is related to the bacterial secretion system and triggers the Hypersensitive Response (HR). The hypersensitive response is the programmed cell death (PCD) of the plant tissue around the site of infection. These lesions are responsible to prevent the spread of the pathogen to other tissues of the plant, thus stopping the development of the disease. As this repressed pathway, the bacterium can reduce the plant's response against the pathogen and favor its spread through the different plant tissues.

Other genes as *SORBIDRAFT_04g005520* (Calmodulin like protein) - repressed 2.9 x and *SORBIDRAFT_08g001200* (calcium-binding protein CML) - repressed 30x have also been identified as differentially expressed, they have the DNA-binding WRKY domain found in transcription factors related to systems such as defense against pathogens, senescence and development of trichomes. These pathways are associated with the innate immune system, being activated by the presence of PAMP's (Pathogen-associated molecular pattern). According to these data, the bacterium is able to modulate the immune response of the plant when in association, decreasing the hypersensitive response and increasing the scattering power in the tissue.

DISCUSSION

Herbaspirillum rubrisubalbicans M1 is a world wide spread bacteria, being found in several different countries. Interestingly the screening of the sorghum genotypes showed high degree of susceptibility in sorghum USA cultivars (54 susceptible and 4 resistant), even from different populations. This fact contrasts with the observed and reported by Pimentel (1991) and Olivares (1997) where this disease was not found in the Brazilian fields and agronomically important cultivars showed resistance, even after the artificial inoculation, which also can be related to the large use of inoculants in the Brazilian agriculture, which does not happen on the United States agriculture. The *Herbaspirillum* genus can be found in several agronomic Brazilian products and this may expose the fields to the presence of this bacterial group in the soil/plants, leading the resistant cultivars to be constantly selected, which does not occur in US fields, even though this bacteria also can be isolated from this country (CHRISTOPHER and EDGERTON, 1930). Recently *H. rubrisubalbicans* was reported for the first time in China causing the Mottled stripe disease in sugarcane (TAN et al., 2010).

James (1997), observed that the bacteria colonize the vascular system, leading to the blockage of these vessels and the red stripe symptoms are probably due to the massive colonization. The necrosis points can be related to the deprivation of nutrients and water caused by the blockage of the vessels.

The pretreatment with the PAMPs indicate that the bacteria can modulate the response against the pathogen. *Herbaspirillum rubrisubalbicans* has a type 3 secretion system (T3SS), this apparatus are found in a range of animal and plant pathogens, but also can be found in the symbiotic organisms (HUECK et al., 1998). This apparatus is known by deliver effector protein inside the host cell and modulate the immune response. Probably the bacteria use this mechanism to delay the plant response and have enough time to colonize and develop the red stripe symptoms. After the pretreatment with flg22 and chitin, the plant was able to respond against the bacteria faster and control the bacterial colonization. This faster response can be related to the callose deposition.

Callose is a β-1,3-Glucan polysaccharide found in association with the cell wall. Their deposition is induced by multiple stresses in higher plants through the activation of the callose synthase enzyme. This enzyme activation can start trough the modification of the cytosolic Ca^{+2} levels and induction of the lipid peroxidation (CHEN and KIM, 2009; KAUSS et al., 1983; FREDRIKSON and LARSSON, 1989; SIVAGURU et al., 2005; ZAVALIEV et al.,

2011). The deposition and accumulation of callose is also related to the β-1,3-glucanases, these enzymes can degradate the callose as well as can degrade the polysaccharide of cell walls and cell surface of pathogens (pathogenesis related protein) (DONG et al. 1991; ZAVALIEV et al., 2011).

Based on our results, 24 hours after bacterial inoculation the plant starts the response by isolated points of callose deposition. When we pretreat this plants with flg22 and chitin we can see that the callose deposition is more pronunciated, with several points of callose deposition homogeneously spread throughout the tissue. Probably, this is the mechanism responsible for the control of the number of bacteria and for the spread of the bacteria inside the tissues. The earlier activation of the immune system can increase the resistance against the pathogen.

Plants are in contact with microbes all the time, pathogenic microbes has the ability to access the internal tissues of the plant by penetrating the leaves and roots, by wounds or by natural openings such as stomata or pores (CHISHOLM et al., 2006). In the *Herbaspirillum rubrisubalbicans* association with sorghum we can see that the stomata is not an important step of the colonization, even with the stomata closed the colonization is successful and the bacteria can causes the disease. This support the idea that *Herbaspirillum* is an endophytic bacteria that can penetrate the tissues by others epidermic discontinuity, can move inside the leaf, and does not need constant re-colonization of epiphytic bacteria.

The interaction between the host and the pathogen is close and one of the ways that this occurs is the secretion by the pathogens of effector proteins, which can modulate the host physiology response against the pathogen. This effector proteins turn the plant more susceptible to the infection (HOGENHOUT et al., 2009; JONES et al., 2006). The plant cell surface has receptors to detected the MAMP`s, these receptors are known as pattern-recognition receptor (PRRs) and can recognize the MAMPs and start the PRR-mediate immune response by the activation of their intracellular kinase domain. Many effectors can act suppressing the PRR-mediated immunity which can improve the capacity of infection of the host (JONES et al., 2006).

The QTL analysis shows a region in the chromosome 7 that is highly related to the Immune response. The function of several different proteins coding by the genes in this location can be involved with the resistance to the red stripe disease. We found different Calcium binding proteins, involved with the ROS response which can modulate the hypersensitive response (a form of programmed cell death), the cell wall reinforcement and the stomata closure. The RNAseq analysis identify the same Calcium binding protein 30 times

down-regulated, suggesting that this protein are closely related to the response against *H. rubrisubalbicans* M1. The CML (calmodulin like proteins) are poor characterized in plants. It is knowing that CML9 in *Arabidopsis* has function as a positive regulator of defense against microbial pathogens and as a negative regulator against abiotic stress (MAGNAN et al. 2008; LEBA et al. 2012). Zhu (2016) shows that the CML8 from *Arabidopsis thaliana* has the expression induced by the inoculation with *Pseudomonas syringae*, acting against the microbial. In contrast, our results shows that the Sorghum CML is Down-regulated after the inoculation with *H. rubrisubalbicans* suggesting that a new pathway of regulation during the pathogen association can be involved with this interaction.

We also found proteins related to the Bacterial T3SS -Type III Secretion System response (PtiI and RIN4) in the QTL and HSP90 in the RNAseq, which support the already described data from SCHIMIDT, 2012, where the bacterial T3SS was described as crucial to the development of the mottle stripe disease in sugarcane, a bacteria mutant in a T3SS component was incapable of causes the disease. Our data confirm the literature, for the first time, by the plant side.

The QTL analysis indicates the presence of *cerk1* gene close to the region of the Chromosome 7 involved in the resistance/susceptibility of sorghum to the red stripe disease. The CERK1 is Lys-M containing Chitin Elicitor Receptor kinase 1, an important receptor for the PAMP perception described in *Arabidopsis thaliana* (LIU et al., 2012; WILLMANN et al., 2012). This receptor has an important well described role in the fungi recognition, but also seems to have a role in the peptidoglycan recognition. Peptidoglycans (PGNs) compound the bacterial cell wall of gram positive and gram-negative cells as β(1,4)-linked N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues (GLAUNER et al., 1988, SCHLEIFER et al., 1972). This polymeric heteroglycans are linked by oligopeptides such as Lysine (Lys) and DAP (meso-diaminopimelic acid) residues, forming a bacterial specific PGNs Lys or DAP-type. Both subtypes of those bacterial PGNs are PAMPs that can trigger the antibacterial immune response in plant (WILLMANN et al., 2011). According to Gimenez-Ibanesz (2009), the lacking of CERK1 in Arabidopsis affects the plant response against the bacterial infection. Willmann (2011) shows that mutants *cerk1-2* in Arabidopsis does not show response neither to fungal chitin or bacteria PGN and also shows that the PGN-induced transcriptome alterations were highly reduced in the *cerk1-2* mutants. These data provide a strong evidence for the participation of CERK1 in PGN recognition and immunity to bacterial infection (WILLMANN et al., 2011). Based on the literature data and in our results we can conclude that CERK1 are involved in the recognition of *Herbaspirillum*

rubrisubalbicans and can modulate the perception and response against this pathogen. A Schematic model of the Sorghum proteins involved in the interaction with *Herbaspirillum rubrisubalbicans* is shown in the Figure 4-16.

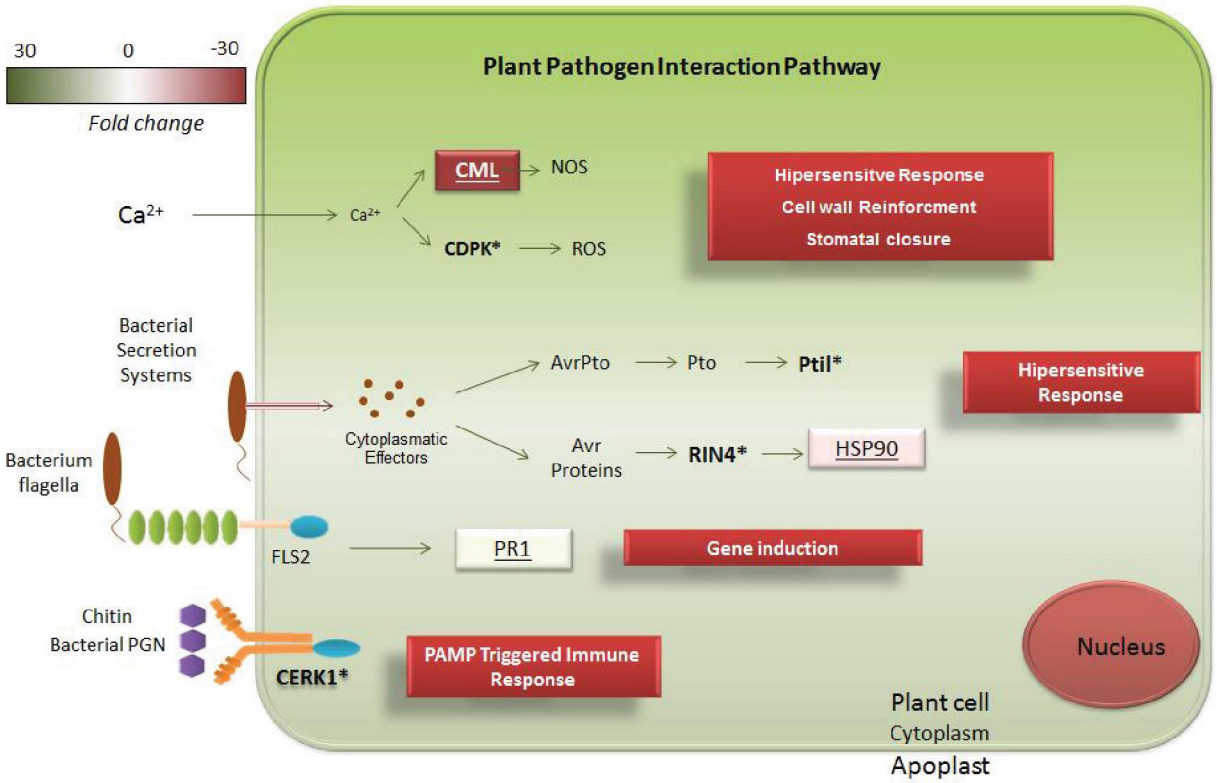


FIGURE 4-16 – Genes related to the Plant Pathogen Interaction identify during the QTL and RNAseq analysis are indicated in red. The genes present in the position 70.81 of Chromosome 7 are indicating in bold and the genes identify during the RNASeq are underlined.

CONCLUSION

Herbaspirillum rubrisubalbicans is an important diazotrophic bacteria which has as interesting pattern of interaction with *Sorghum bicolor*. In this work we show that many factors can be involved with this interaction. Many classic pathways are involved and, since this bacteria is a mild pathogen and it is largely used in the agriculture, this system can be a great option for the Sorghum immune system study. Sorghum is a crop with biofuel use perspectives and the total understanding of their interactions with microbes is crucial for the total use of this plant in all capacities.

EXPERIMENTAL PROCEDURES

Bacterial Strain and Growth conditions

The *Herbaspirillum rubrisubalbicans* M1 was growth in NFb-HP Malate media with 20mM of NH4Cl as nitrogen source, 30°C, 120 rpm for 16-24 hours or until reach the OD_{600} of 1. For the inoculation, the culture were washed 3 times with saline solution (0.9% NaCl) to remove all the NFbHP malate media.

Sorghum germination and inoculation.

The sorghum seeds were planted directly into the humid soil (Sunshine) and kept covered by plastic lid in a microenvironment for 5 days until the completely germination. The plants were kept in green house with the temperature control in approximately 23-25°C and were watered every 3 days with non-sterile water.

After 7 days, the plants were inoculated with 500 µL of saline solution (NaCl 0.9%) containing 10^8 bacterial cells per mL (OD_{600}=1). The inoculation was performed by the injection of the solution on the stalk of the plants, approximately 1 cm above the soil with a hypodermic needle. The plants were observed for more 5-10 days or until the development of the disease.

Scanning Electron Microscopy

The samples were fixed in Karnovisky solution (2% paraformaldehyde, 2.5% glutaraldehyde in 0.1M cacodylate buffer pH 7.2 at 4° C) (KARNOVSKY et al., 1965), dehydrated in ascending alcohol and acetone series, and the critical point was reached in a Bal-Tec CPD - 030 with carbon dioxide. The plating in gold was obtained in a Balzers SCD - 030. SEM analysis were performed in a scanning electron microscope JEOL-JSM 6360 LV.

PAMP`s assay

The sorghum plants were growth for 7 days in greenhouse and were inoculated by infiltration using a syringe with approximately 100 µL of 100 nM of flg22 and 1µM of chitin. After 2 hours the same location was infiltrated with a 100 µL of bacteria solution, containing 10^8 cells per mL in NaCl 0.9%. For the control, the plant was inoculated with saline solution (NaCl 0.9%) and 2 hours later the bacteria was inoculated. The plants were observed for 7-10 days until the development of the symptoms.

Image analysis and measurement of lesions

The measurement of the lesions were perform in the Matlab program using the K-means algorithm. The algorithm was programmed for split the picture in 3 different groups: Background, Leaf and Symptom. After the first step the background image was recovered, the borders of the leaves were remodel and the necrosis points were extracted and counted as a symptom area. Then the symptom area was calculated based on the necrotic areas and in the red stripes disease. The ratio between the control leaves symptom area and the treated leaves symptom area was used for the analysis.

Count of bacteria inside the tissues

For the count of bacteria in different portions of the leaf, we split the leaves in the interesting portions and cut the pieces with a puncher resulting in circles of 0.5 – 0.7 cm of diameter. Then the pieces of leaves were washed 3 times with Saline Solution and macerated in the same solution. The solution was serial diluted and plated in NFbHP Malate with 20mM of NH_4Cl. The plates were kept in growth chamber – 30C.

To analyze the spread of the bacteria in the leaf, successive portions with 1 cm of distance were punched and counted in the same way as described above.

Callose identification

The callose quantification was made using Aniline blue. The plants were recovered after the treatments and place in a 50 mL tube with Ethanol 96% until total chlorophyll removal (1-2 days). The leaves were rehydrated in phosphate solution (5.2 g of K_2HPO_4 in 200 ml H_2O) for 30 minutes. Then the leaves were staining with 0.01% of aniline blue solution for at least 2 hours. The leaves were washed 3 times with phosphate solution and observed in fluorescent microscopy with DAPI filter (345nm of excitation).

ABA Assay

The plants were growth for 7 days, and were inoculated by infiltration with 100 µL of Abscissic Acid (ABA). The stomata status was checked before and after the infiltration with ABA with light microscopy. The bacteria was infiltrate in the same location (200 µL of a 10^8 bacteria/mL solution). The plants were observed for 7 days after inoculation, until the symptoms development.

QTL

Plant Materials. To map flg22 and chitin-related loci, a $F_{4:7}$ biparental mapping population consisting of 103 recombinant inbred lines (RILs) was selected based on the differential response between a low-responding parent, BTx623, and a high-responding parent, SC155-14E. The sorghum seeds were planted directly into the humid soil (Sunshine) and kept covered by plastic lid in a microenvironment for 5 days until the completely germination. The plants were kept in green house with the temperature control in approximately 23-25°C and were watered every 3 days with non-sterile water.

Genotypic Analysis & Linkage Mapping. A genetic linkage map consisting of 956 polymorphic single nucleotide polymorphisms (SNPs) was obtained from Dr. Patricia Klein at Texas A&M University (Patil et al., 2017).

Disease Testing. For each experimental run, plants were arranged in a randomized complete block design with three replications. Three experimental runs were evaluated. After 7 days, the plants were inoculated with 500 µL of saline solution (NaCl 0.9%) containing 10^8 bacterial cells per mL (OD_{600}=1). The inoculation was performed by the injection of the solution on the stalk of the plants, approximately 1 cm above the soil with a hypodermic needle. The plants were observed for more 5-10 days or until the development of the disease.

Data Collection. Two rating systems were utilized to score disease severity. A visual 0 to 5 pathogenicity scale, where 0 means no symptom and 5 means high levels of symptoms, was used (supplementary figure 3). For the count of bacteria in different portions of the leaf, we split the leaves in the interesting portions and cut the pieces with a puncher resulting in circles of 0.5 – 0.7 cm of diameter. Then the pieces of leaves were washed 3 times with Saline Solution and macerated in the same solution. The solution was serial diluted and plated in NFbHP Malate with 20mM of NH4Cl. The plates were kept in growth chamber – 30C. To analyze the spread of the bacteria in the leaf, successive portions with 1 cm of distance were punched and counted in the same way as described above.

Phenotypic Data Analysis. The normality of data was checked using PROC UNIVARIATE in SAS statistical software version 9.4 (SAS Institute, 2015) and subsequently, bacterial counts were log-transformed to achieve normality. An analysis of variance (ANOVA) and least square (LS) means were generated using the GLM procedure in SAS statistical software version 9.4 (SAS Institute, 2015) for both visual disease ratings and bacterial counts. Standard F-tests in all analyses were used to determine significance of main effects and interactions. In the model, replications and experimental runs were considered random. Individual hypothesis testing of effects and interactions was designated using an

appropriate error term in the GLM procedure. Pearson correlation coefficients were calculated to compare both replications and experimental runs for visual disease ratings and bacterial counts.

QTL Mapping. Using Windows QTL cartographer 2.5 (Wang et al., 2011), composite interval mapping (CIM) was used to identify QTL, estimate QTL LOD scores, and measure the proportion of phenotypic variation explained by individual QTL. Default settings for CIM were used, and a permutation test was conducted to simulate estimates for unknown parameters and establish thresholds for the detection of putative QTL (Churchill and Doerge, 1994).

RNAseq

The plants were grown in vermiculate for 14 days, then were inoculated with 107 bactereia using a hypodermic syringe in the stalk. Seven days after inoculation the leaves were harvested, tranfered directly to liquid nitrogen and, macerated in RNA later solution.

The RNA was extracted using Trizol (ThermoFisher Scientific) and treated with DNaseI (ThermoFisher Scientific) according to manufacturer instructions. The rRNA was depleted and the librarys were constructed following the manufacturer instructions. The librarys were sequenced in Next Generation sequencing platform **Ion Proton™ System** - Life Technologies. Each treatment were prepared in biological triplicate, and each replicate was sequenced at least 3 times. The mapping was performed using the CLC Workebench Softwere – version 7.1 and the differential expression genes were obtained using R program, packages DESeq and EdgeR.

ACKNOWLEDGMENTS

We are grateful to Roseli Prado (UFPR), Valter A. Baura (UFPR) and Yaya Cui (University of Missouri) for their technical support, to Electron Microscope Center - Federal University of Parana and Electron *Microscopy* Core *Facility* (EMC) – University of Missouri for the microscope images. This work was supported by CNPq/INCT, CAPES and Fundação Araucária.

AUTHOR CONTRIBUTION

Jennifer Kimball (Department of Plant Pathology, North Carolina State University, USA) – Assisted the QTL analysis;

Fernanda P do Amaral and Beverly Jose Agtuca (Division of Plant Science and Biochemistry, University of Missouri, USA) - Assisted the Sorghum Screening;

Tomas P Pereira (Department of Science and Food Technology, Federal University of Santa Catarina, Brazil) - Assisted the Sorghum Screening;

Gabriel H. A. Pereira (Department of Geodetic Sciences, Federal University of Paraná, Brazil) – Asssited the image analysis;

Michelle Zibetti Tadra Sfeir (Department of Biochemistry and Molecular Biology, Federal University of Paraná, Brazil) – Assited the RNASeq assay;

Fabio de Oliveira Pedrosa, Emanuel Maltempi de Souza and Rose Adele Monteiro (Department of Biochemistry and Molecular Biology, Federal University of Paraná, Brazil) – Orientation and Supervision;

Peter Balint-Kurti (Department of Plant Pathology, North Carolina State University, USA) - Orientation and Supervision;

Gary Stacey (Division of Plant Science and Biochemistry, University of Missouri, USA) - Orientation and Supervision.

REFERENCES

BALDANI, VLD., BALDANI, J. I., OLIVARES, F., DOBEREINER, J. Identification and ecology of *Herbaspirillum* seropedicae and the closely related Pseudomonas rubrisubalbicans. **Symbiosis** 1992.

BALDANI, J.I, POT, B., IRCHHOF, G., ALSEN, E., BALDANI, V. L. D., OLIVARES, F. L., HOSTE, B., KERSTERS, K., HARTMANN, A., GILLIS, M., DÖBEREINER, J. Emended Description of *Herbaspirillum*; Inccusion of [Pseudomonas] rubrisubalbicans, a Mild Plant Pathogen, as *Herbaspirillum* rubrisubalbicans comb. nov.; and Classification of a Group of Clinical Isolates (EF Group 1) as *Herbaspirillum* Species 3. **International Journal of Systematic and Evolutionary Microbiology**, 1996

BODDEY RM, DE OLIVEIRA OC, URQUIAGA S, REIS VM, OLIVARES FL, BALDANI VLD, DÖBEREINER J (1995) Biological nitrogen fixation associated with sugar cane and rice: contributions and prospects for improvement. **Plant Soil** 174:195–209

BOYER JS, Plant productivity and environment. **Science** 218:443–448 (1982).

CHRISTOPHER, W. N., EDGERTON, C. W Bacterial stripe diseases of Sugar Cane in Louisiana. **Journal of Agricultural Research**. 1930.

CRUZ LM, SOUZA EM, WEBER OB, BALDANI JI, DÖBEREINER J, PEDROSA FO. 16S Ribosomal DNA characterization of nitrogen-fixing Bacteria isolated from banana (Musa spp.) and pineapple (Ananas comosus (L.) Merril). **Appl Environ Microbiol** 67:2375–2379. 2001

ELBELTAGY A, NISHIOKA K, SATO T Endophytic colonization and in plant nitrogen fixation by a *Herbaspirillum* sp. isolated from wild rice species. **Appl Environ Microbiol** 67:5285–5293, 2001.

GALLI F, CARVALHO PCT, TOKESHI H, BALMER F, KIMATI H, CARDOSO CO, SALGASO CL, KRUGNER TL, CARDOSO EJBN, BERGAMIN FA Manual de fitopatologia: doenças de plantas cultivadas. **Agronômica Ceres**, São Paulo, 1980.

GLAUNER B, HÖLTJE JV, SCHWARZ U The composition of the murein of Escherichia coli. **J Biol Chem** 263:10088–10095. 12. (1988)

GYANESHWAR P, JAMES EK, REDDY PM, LADHA JK (2002) *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. **New Phytol** 154:131–145

HALE, C. N., AND J. P. WILKE. 1972. A comparative study of Pseudomonas species pathogenic to sorghum. **N. Z. J. Agric. Res**. 15:448–456.

HOGENHOUT, S. A., VAN DER HOORN, R. A., TERAUCHI, R. & KAMOUN, S. Emerging concepts in effector biology of plant-associated organisms. **Mol. Plant Microbe Interact.** 22, 115–122 (2009).

JAMES, E. K., OLIVARES, F. L., BALDANI, J. I., DÖBEREINER, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue 3Sorghum bicolor L. Moench **Journal of Experimental Botany** 48 (3), 785-798. 1997

JAMES, E. K., OLIVARES, F. L. Infection and colonization of sugar cane and other graminaceous plants by endophytic diazotrophs. **Critical Reviews in Plant Sciences** 17 (1), 77-119 1998

JAMES EK, GYANESHWAR G, BARRAQUIO WL, LADHA JK Endophytic diazotrophs associated with rice. In: Ladha JK, Reddy PN (eds) The quest for nitrogen fixation in rice. **International Rice Research Institute**, Manila (1999)

JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Research**. Volume 65, Issues 2–3, March, Pages 197–209. 2000

JAMES EK, GYANESHWAR P, MATHAN N Infection and colonization of rice seedlings by the plant growthpromotion bacterium *Herbaspirillum* seropedicae Z67. **Mol Plant Microbe Interact** 15:894–906 (2002)

JONES, J. D., DANGL, J. L. The plant immune system. **Nature** 444, 323. 2006.

LEBA, L.J., CHEVAL, C., ORTIZ-MARTÍN, I., RANTY, B., BEUZÓN, C.R., GALAUD, J.P., et al. CML9, an Arabidopsis calmodulin-like protein, contributes to plant innate immunity through a flagellin-dependent signalling pathway. **Plant J.** 71: 976-989. 2012

LIU, T. et al. Chitin-induced dimerization activates a plant immune receptor. **Science** 336, 1160–1164 (2012).

MAGNAN, F., RANTY, B., CHARPENTEAU, M., SOTTA, B., GALAUD, J.P. AND ALDON, D. Mutations in AtCML9, a calmodulin-like protein from Arabidopsis thaliana, alter plant responses to abiotic stress and abscisic acid. **Plant J**.56: 575-589. (2008)

MCBEE, GG, MILLER FR, DOMINY RE AND MONK RL, Quality of sorghum biomass for methanogenesis. **Energy from Biomass and Waste X**, ed. by Klass DL. Elsevier, London. pp. 251–260. 1987.

OLIVARES, F.L., JAMES, E. K., BALDANI, J. I., DÖBEREINER, J. Infection of mottled stripe disease-susceptible and resistant sugar cane varieties by the endophytic diazotroph Herbaspirilium. **New Phytologist** 135 (4), 723-737. 1997.

OLIVEIRA ALM, CANUTO EL, URQUIAGA S, REIS VM, BALDANI JI: Yield of micropropagated sugarcane varieties in different soil types following inoculation with diazotrophic bacteria**. Plant Soil**., 284: 23-32. 2006.

OLIVEIRA ALM, STOFFELS M, SCHMID M, REIS VM, BALDANI JI, HARTMANN A: Colonization of sugarcane plantlets by mixed inoculations with diazotrophic bacteria. **Eur J Soil Biol**. 2009.

PIMENTEL, J.P., OLIVARES, F., PITARD, R. M., URQUIAGA, S., AKIBA, F., DOBEREINER, J. Dinitrogen fixation and infection of grass leaves by *Pseudomonas rubrisubalbicans* and *Herbaspirillum seropedicae*. **Plant and Soil**, 1991

RONCATO-MACCARI LDB, RAMOS HJO, PEDROSA FO, ALQUINI Y, CHUBATSU LS, YATES MG, RIGO LU, STEFFENS MBR, SOUZA EM Root colonization, systemic spreading and contribution of *Herbaspirillum* seropedicae to growth of rice seedlings. **Symbiosis** 35:01–10. 2003.

ROONEY, W. L., BLUMENTHAL, J., BEAN, B. AND MULLET, J. E. Designing sorghum as a dedicated bioenergy feedstock. **Biofuels**, Bioprod. Bioref., 1: 147–157. 2007.

SCHLEIFER KH, KANDLER O Peptidoglycan types of bacterial cell walls and their taxonomic implications. **Bacteriol Rev** 36:407–477. 1972.

TAN, Z. Q., MEN, R., ZHANG, R. Y., HUANG, Z. First Report of *Herbaspirillum rubrisubalbicans* Causing Mottled Stripe Disease on Sugarcane in China. **Plant disease - Disease Notes**. March 2010, Volume 94, Number 3

XIAOYANG ZHU, X., ROBE, E., JOMAT, L., ALDON, D., MAZARS, C., GALAUD, J.P. CML8, an Arabidopsis calmodulin-like protein plays a role in Pseudomonas syringae plant immunity**. Plant Cell Physiology**. 2016.

WIEDENFELD, R.P. Nutrient requirements and use efficiency by sweet sorghum. **Energy in Agriculture**, 3 49--59 49. 1984.

WILLMANN, R. & NURNBERGER, T. How plant lysin motif receptors get activated: Lessons learned from structural biology. **Sci. Signal**. 5, e28. 2012.

SUPPLEMENTARY TABLE 4-3 – Genes Up and Down Regulated in Sorhum leaves inoculated with *Herbaspirillum rubrisubalbicans* M1 identify using DESeq package (R program) during the RNASeq analysis.

UP Regulated Genes				
Gene Id	**Annotation**	**Probable function**	**Fold Change**	***p value***
SORBIDRAFT_01g008910	hypothetical protein	beta-fructofuranosidase	179,97	4,35E-21
SORBIDRAFT_07g028620	hypothetical protein	probable galactinol--sucrose galactosyltransferase 1	134,01	2,20E-04
SORBIDRAFT_08g021500	hypothetical protein	Domain of unknown function	101,37	9,67E-07
SORBIDRAFT_01g040050	hypothetical protein	The first Lim domain of pollen specific protein SF3	91,42	5,27E-05
SORBIDRAFT_09g024060	hypothetical protein	Glycogen recognition site of AMP-activated protein kinase	85,92	1,18E-04
SORBIDRAFT_03g012340	hypothetical protein	IGR protein motif	76,59	3,65E-04
SORBIDRAFT_02g029460	hypothetical protein	No apical meristem (NAM) protein; pfam02365	68,11	3,75E-25
SORBIDRAFT_01g044130	hypothetical protein	L,L-diaminopimelate aminotransferase	67,39	1,16E-03
SORBIDRAFT_02g007240	hypothetical protein	protein SRG1	47,93	8,20E-09
SORBIDRAFT_02g038680	hypothetical protein	similar to Putative peptide transporter	44,93	5,05E-08
SORBIDRAFT_02g003590	hypothetical protein	Transferase family	37,13	1,51E-04
SORBIDRAFT_01g041360	hypothetical protein		35,51	1,41E-03
SORBIDRAFT_04g027570	hypothetical protein	similar to Oxidation protection protein-like	32,76	5,17E-62
SORBIDRAFT_08g011530	hypothetical protein	laccase	31,52	3,04E-04
SORBIDRAFT_02g043180	hypothetical protein	Poly-adenylate binding protein, unique domain	29,44	7,33E-05
SORBIDRAFT_06g023450	hypothetical protein	Glycosyl hydrolase family 3 N terminal domain	29,36	5,72E-05
SORBIDRAFT_09g001210	hypothetical protein	Glycoside hydrolase family 19 chitinase domain	29,24	1,31E-07
SORBIDRAFT_06g028360	hypothetical protein	DNA binding site [nucleotide binding]	29,21	4,05E-07
SORBIDRAFT_03g002320	hypothetical protein	Phytochelatin synthase	26,08	1,34E-07
SORBIDRAFT_06g020720	hypothetical protein	malate synthase	25,51	2,10E-12
SORBIDRAFT_03g041770	hypothetical protein	leucine-rich repeat [structural motif]	24,31	3,87E-05
SORBIDRAFT_05g001690	hypothetical protein	ZF-HD protein dimerization region	24,15	4,85E-05
SORBIDRAFT_04g025930	hypothetical protein	adenine phosphoribosyltransferase	22,80	1,10E-04
SORBIDRAFT_02g030050	hypothetical protein	UDP:flavonoid glycosyltransferase YjiC, YdhE	20,15	2,86E-06

Continuation

		family		
SORBIDRAFT_06g016530	hypothetical protein	ABA/WDS induced protein	19,31	1,98E-07
SORBIDRAFT_01g033660	hypothetical protein		19,18	3,93E-05
SORBIDRAFT_01g034160	hypothetical protein	Amino acid transporter	18,49	7,94E-06
SORBIDRAFT_05g022940	hypothetical protein	Barwin family similar to Pathogenesis-related protein 4	17,18	2,20E-16
SORBIDRAFT_04g003140	hypothetical protein	DNA-binding domain	16,97	1,51E-05
SORBIDRAFT_03g045710	hypothetical protein		16,92	1,74E-05
SORBIDRAFT_02g030540	hypothetical protein	Rab subfamily motif 1 (RabSF1	16,34	5,21E-09
SORBIDRAFT_01g037960	hypothetical protein		15,53	8,93E-13
SORBIDRAFT_06g029550	hypothetical protein	anthocyanidin reductase	14,30	4,91E-18
SORBIDRAFT_08g022440	hypothetical protein		13,63	8,15E-07
SORBIDRAFT_09g029300	hypothetical protein	0	13,52	1,64E-06
SORBIDRAFT_03g044780	hypothetical protein	K+ potassium transporter	13,13	5,37E-05
SORBIDRAFT_01g046360	hypothetical protein	similar to 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	11,51	2,30E-04
SORBIDRAFT_10g023140	hypothetical protein		11,38	2,87E-10
SORBIDRAFT_05g006950	hypothetical protein		11,23	5,04E-06
SORBIDRAFT_03g005920	hypothetical protein	oxidoreductase, 2OG-Fe(II) oxygenase family	11,17	6,56E-05
SORBIDRAFT_02g003580	hypothetical protein	leucine-rich repeat receptor-like protein kinase	11,13	2,86E-07
SORBIDRAFT_03g012910	hypothetical protein	2-oxoglutarate dehydrogenase (OGDH)	10,93	4,52E-04
SORBIDRAFT_07g025770	hypothetical protein	tress up-regulated Nod 19	10,93	2,85E-08
SORBIDRAFT_01g026980	hypothetical protein	redicted oxidoreductase	10,91	6,90E-09
SORBIDRAFT_06g025650	hypothetical protein	bHLH-MYC and R2R3-MYB transcription factors	10,51	7,62E-14
SORBIDRAFT_04g027860	hypothetical protein	he first cupredoxin domain of plant laccases	10,46	3,37E-05
SORBIDRAFT_02g030740	hypothetical protein		10,42	1,29E-04
SORBIDRAFT_08g020020	hypothetical protein		10,25	1,70E-06
SORBIDRAFT_10g004580	hypothetical protein	DNA-binding domain	10,23	2,89E-06
SORBIDRAFT_05g018800	hypothetical protein	Leucine rich repeat N-terminal domain	10,18	1,82E-20
SORBIDRAFT_02g003090	hypothetical protein	Thioredoxin_like	10,10	1,74E-06
SORBIDRAFT_01g042690	hypothetical protein		10,07	6,36E-08
SORBIDRAFT_04g013083	hypothetical protein	"5'-3' exonuclease	9,90	6,83E-04
SORBIDRAFT_05g002710	hypothetical protein	Glycosyltransferase_GTB_type	9,17	4,82E-04
SORBIDRAFT_05g006880	hypothetical protein		9,03	1,05E-14
SORBIDRAFT_01g012710	hypothetical protein	glycoside hydrolase family 64	8,95	5,71E-11
SORBIDRAFT_09g022360	hypothetical protein	Aldo/keto reductase	8,85	3,69E-07
SORBIDRAFT_06g022660	hypothetical protein	Myb-like DNA-binding domain	8,74	1,15E-05

Continuation

SORBIDRAFT_05g020220	hypothetical protein	chalcone synthase	8,70	1,34E-05
SORBIDRAFT_09g005910	hypothetical protein		8,66	9,62E-11
SORBIDRAFT_08g016200	hypothetical protein	Cytochrome P450	8,62	6,17E-04
SORBIDRAFT_01g046960	hypothetical protein	EamA-like transporter family	8,57	7,58E-05
SORBIDRAFT_06g000500	hypothetical protein	The Major Facilitator Superfamily (MFS)	8,45	2,79E-06
SORBIDRAFT_01g036650	hypothetical protein	aminophospholipid translocase	8,28	2,27E-05
SORBIDRAFT_03g027480	hypothetical protein	ABC transporter G family member	8,24	4,90E-08
SORBIDRAFT_09g026090	hypothetical protein	leucine-rich repeat receptor-like protein kinase	8,16	1,27E-05
SORBIDRAFT_06g025250	hypothetical protein	Prolyl oligopeptidase PreP, S9A serine peptidase family	8,05	1,29E-20
SORBIDRAFT_02g007460	hypothetical protein	nucleoside_deaminase	8,03	1,10E-07
SORBIDRAFT_02g037570	hypothetical protein	Sugar (and other) transporter	7,84	1,46E-06
SORBIDRAFT_01g045940	hypothetical protein		7,44	1,69E-05
SORBIDRAFT_02g002150	hypothetical protein	SCP_PR-1_like: SCP-like extracellular protein domain, PR-1 like subfamily. The wider family of SCP containing proteins includes plant pathogenesis-related protein 1 pathogens, and may act as an anti-fungal(PR-1), which accumulates after infections	7,33	2,02E-17
SORBIDRAFT_06g021250	hypothetical protein	Chitin recognition protein	7,10	1,48E-08
SORBIDRAFT_08g020880	hypothetical protein	Wall-associated receptor kinase galacturonan-binding	7,05	9,67E-04
SORBIDRAFT_05g020160	hypothetical protein	Chalcone and stilbene synthases	6,95	1,00E-06
SORBIDRAFT_01g013080	hypothetical protein	short chain dehydrogenase	6,89	1,07E-09
SORBIDRAFT_04g001130	hypothetical protein	catalase_clade_1	6,61	7,85E-24
SORBIDRAFT_04g028610	hypothetical protein	oxidoreductase	6,59	1,61E-09
SORBIDRAFT_05g020200	hypothetical protein	chalcone synthase	6,43	3,06E-06
SORBIDRAFT_09g015910	hypothetical protein	Glyoxysomal and mitochondrial malate hydrogenase	6,29	1,07E-03
SORBIDRAFT_05g019070	hypothetical protein	Rab interacting lysosomal protein-like 1 and 2	6,23	1,06E-08
SORBIDRAFT_03g011745	hypothetical protein	Isopenicillin N synthase and related dioxygenases	6,22	7,17E-07
SORBIDRAFT_09g001180	hypothetical protein		5,93	1,67E-06
SORBIDRAFT_03g010040	hypothetical protein	5'-3' exonuclease	5,88	1,77E-09
SORBIDRAFT_08g006710	hypothetical protein	Leucine-rich repeat (LRR) protein	5,86	2,82E-04

Continuation

SORBIDRAFT_03g045460	hypothetical protein	Glycosyl hydrolases family 17	5,80	1,03E-17
SORBIDRAFT_03g037040	hypothetical protein	3-hydroxybutyryl-CoA dehydrogenase	5,78	1,67E-03
SORBIDRAFT_03g024830	hypothetical protein	methyl indole-3-acetate methyltransferase	5,76	2,67E-09
SORBIDRAFT_03g029290	hypothetical protein		5,74	2,75E-06
SORBIDRAFT_04g021780	hypothetical protein		5,67	1,89E-05
SORBIDRAFT_03g026960	hypothetical protein	Calcium-binding EGF-like domain	5,64	4,65E-04
SORBIDRAFT_03g044830	hypothetical protein		5,55	5,92E-07
SORBIDRAFT_07g025520	hypothetical protein		5,55	2,90E-11
SORBIDRAFT_10g028690	hypothetical protein		5,40	1,58E-05
SORBIDRAFT_01g011890	hypothetical protein	Nodulin-like	5,38	7,06E-07
SORBIDRAFT_05g026870	hypothetical protein		5,33	7,13E-05
SORBIDRAFT_09g021800	hypothetical protein	Glycosyl hydrolase family 1	5,28	1,09E-03
SORBIDRAFT_01g045980	hypothetical protein		5,27	3,02E-05
SORBIDRAFT_01g031000	hypothetical protein	Glutathione S-transferase	5,24	1,29E-05
SORBIDRAFT_05g024090	hypothetical protein	Leucine rich repeat N-terminal domain	5,22	4,35E-23
SORBIDRAFT_09g005800	hypothetical protein	Hpt domain	5,21	1,78E-03
SORBIDRAFT_08g001430	hypothetical protein	Protein Kinases,	5,16	2,20E-04
SORBIDRAFT_07g020270	hypothetical protein	Trehalose-6-Phosphate Synthase (TPS)	5,13	1,06E-06
SORBIDRAFT_10g022190	hypothetical protein	ABC transporter transmembrane region	5,11	7,18E-05
SORBIDRAFT_04g003110	hypothetical protein	Methyltransf_7	4,99	7,96E-09
SORBIDRAFT_09g027950	hypothetical protein	non-haem dioxygenase in morphine synthesis	4,97	3,02E-10
SORBIDRAFT_03g001350	hypothetical protein		4,85	4,34E-12
SORBIDRAFT_10g004950	hypothetical protein	Alpha amylase catalytic domain family	4,84	7,96E-05
SORBIDRAFT_02g009660	hypothetical protein	Membrane-associating domain	4,82	2,79E-10
SORBIDRAFT_06g021930	hypothetical protein	Glycosyltransferase family 29	4,76	6,45E-06
SORBIDRAFT_05g022280	hypothetical protein		4,69	1,71E-06
SORBIDRAFT_02g028160	hypothetical protein		4,60	5,58E-08
SORBIDRAFT_03g042630	hypothetical protein		4,58	6,57E-13
SORBIDRAFT_08g020250	hypothetical protein	allantoate amidohydrolase	4,57	1,59E-05
SORBIDRAFT_02g028510	hypothetical protein	Membrane transport protein	4,56	8,55E-09
SORBIDRAFT_01g020810	hypothetical protein		4,54	8,01E-09
SORBIDRAFT_01g027020	hypothetical protein	Superfamily II DNA helicase RecQ	4,53	9,74E-07
SORBIDRAFT_10g028950	hypothetical protein	Glycosyltransferase family A	4,47	3,19E-05
SORBIDRAFT_05g018520	hypothetical protein	OPT oligopeptide transporter protein	4,47	2,80E-04
SORBIDRAFT_02g031870	hypothetical protein	DNA recombination-mediator	4,45	8,14E-16

Continuation

		protein A		
SORBIDRAFT_06g014930	hypothetical protein	potassium uptake protein	4,45	4,89E-06
SORBIDRAFT_10g000240	hypothetical protein	Methyltransf_FA	4,43	2,31E-09
SORBIDRAFT_10g024980	hypothetical protein		4,42	1,40E-03
SORBIDRAFT_08g022450	hypothetical protein		4,42	6,56E-09
SORBIDRAFT_03g002180	hypothetical protein		4,41	1,07E-03
SORBIDRAFT_08g000260	hypothetical protein	PAS domain	4,38	8,68E-05
SORBIDRAFT_06g022500	hypothetical protein	Glycosyl hydrolase family 1	4,34	1,19E-24
SORBIDRAFT_02g003030	hypothetical protein	Serine/Threonine protein kinases	4,33	2,25E-04
SORBIDRAFT_03g043420	hypothetical protein		4,33	6,47E-22
SORBIDRAFT_04g023530	hypothetical protein	Glycosyltransferase_GTB_type	4,32	5,98E-06
SORBIDRAFT_03g028650	hypothetical protein	Cytochrome P450	4,32	1,05E-04
SORBIDRAFT_09g026410	hypothetical protein	HSP70 interaction site	4,29	6,91E-23
SORBIDRAFT_05g026010	hypothetical protein	leucine-rich repeat receptor-like protein kinase	4,29	2,19E-12
SORBIDRAFT_01g017270	hypothetical protein		4,29	3,44E-05
SORBIDRAFT_04g033720	hypothetical protein		4,22	2,22E-04
SORBIDRAFT_06g026830	hypothetical protein		4,21	9,93E-12
SORBIDRAFT_03g015670	hypothetical protein	NAD(P)+-dependent aldehyde dehydrogenase	4,20	5,21E-10
SORBIDRAFT_04g031400	hypothetical protein	Purine nucleobase transmembrane transport	4,19	4,19E-04
SORBIDRAFT_07g000280	hypothetical protein		4,18	6,80E-11
SORBIDRAFT_01g010940	hypothetical protein		4,16	5,14E-05
SORBIDRAFT_01g017440	hypothetical protein		4,14	1,81E-07
SORBIDRAFT_09g002260	hypothetical protein	Isovaleryl-CoA dehydrogenase	4,14	8,66E-12
SORBIDRAFT_01g013150	hypothetical protein		4,09	1,67E-09
SORBIDRAFT_01g007070	hypothetical protein		4,06	5,79E-07
SORBIDRAFT_03g012950	hypothetical protein	Late embryogenesis abundant protein	4,05	7,47E-09
SORBIDRAFT_09g019650	hypothetical protein	alpha/beta hydrolases	4,00	1,13E-14
SORBIDRAFT_05g023150	hypothetical protein		3,94	4,03E-07
SORBIDRAFT_01g037970	hypothetical protein	hydrophobic ligand binding site	3,90	4,63E-18
SORBIDRAFT_06g025570	hypothetical protein	Polyphenol oxidase middle domain	3,82	6,88E-25
SORBIDRAFT_09g002480	hypothetical protein	acetylornithine transaminase	3,80	1,40E-04
SORBIDRAFT_07g025220	hypothetical protein	L-idonate 5-dehydrogenase	3,78	7,35E-05
SORBIDRAFT_07g005610	hypothetical protein	No apical meristem (NAM) protein	3,76	3,67E-04
SORBIDRAFT_09g001320	hypothetical protein	Metallothio_2	3,63	6,21E-19
SORBIDRAFT_03g039510	hypothetical protein		3,63	1,37E-07
SORBIDRAFT_02g040610	hypothetical protein	Mitochondrial branched-chain alpha-ketoacid dehydrogenase kinase	3,62	3,51E-15

Continuation

SORBIDRAFT_01g036560	hypothetical protein		3,61	3,28E-05
SORBIDRAFT_01g038840	hypothetical protein	Rab-GTPase-TBC domain	3,58	2,22E-04
SORBIDRAFT_01g019380	hypothetical protein	photosystem II biogenesis protein Psb29	3,55	1,46E-08
SORBIDRAFT_07g020100	hypothetical protein		3,49	8,37E-10
SORBIDRAFT_05g026040	hypothetical protein	leucine-rich repeat receptor-like protein kinase	3,47	2,84E-08
SORBIDRAFT_02g025790	hypothetical protein	NAD binding site	3,38	2,80E-07
SORBIDRAFT_01g043400	hypothetical protein	Solute carrier families 5 and 6-like	3,33	1,11E-11
SORBIDRAFT_05g002860	hypothetical protein	Glutamine amidotransferases class-II	3,32	5,39E-06
SORBIDRAFT_07g021090	hypothetical protein	Glycosyltransferase_GTB_ type	3,32	3,69E-06
SORBIDRAFT_05g008800	hypothetical protein		3,31	7,66E-05
SORBIDRAFT_03g004490	hypothetical protein	Universal stress protein family	3,20	4,81E-08
SORBIDRAFT_04g032250	hypothetical protein		3,18	4,09E-04
SORBIDRAFT_03g043990	hypothetical protein		3,16	1,59E-09
SORBIDRAFT_01g022260	hypothetical protein		3,11	4,37E-12
SORBIDRAFT_01g013600	hypothetical protein	4-aminobutyrate aminotransferase	3,10	4,45E-07
SORBIDRAFT_07g005920	hypothetical protein		3,09	1,16E-05
SORBIDRAFT_06g027595	hypothetical protein		3,08	3,68E-04
SORBIDRAFT_04g030010	hypothetical protein	iacylglycerol O-acyltransferase	3,08	5,99E-05
SORBIDRAFT_02g005070	hypothetical protein	P-loop_NTPase	3,08	1,10E-07
SORBIDRAFT_06g023010	hypothetical protein	OPT oligopeptide transporter protein	3,02	2,54E-09
SORBIDRAFT_07g003870	hypothetical protein		2,98	2,53E-08
SORBIDRAFT_09g002470	hypothetical protein	glutamate--cysteine ligase	2,98	8,55E-10
SORBIDRAFT_09g030280	hypothetical protein		2,95	2,15E-07
SORBIDRAFT_08g022770	hypothetical protein	MDH_glyoxysomal_mitochondrial	2,94	4,27E-05
SORBIDRAFT_02g036270	hypothetical protein		2,93	8,12E-10
SORBIDRAFT_04g017830	hypothetical protein	Cathepsin propeptide inhibitor domain	2,92	1,32E-17
SORBIDRAFT_01g047690	hypothetical protein	similar to POT family protein	2,92	1,52E-08
SORBIDRAFT_01g047120	hypothetical protein	pheophorbide a oxygenase	2,92	3,45E-10
SORBIDRAFT_03g025440	hypothetical protein		2,84	3,32E-17
SORBIDRAFT_01g005650	hypothetical protein	aldehyde oxidase	2,83	4,64E-14
SORBIDRAFT_01g013000	hypothetical protein		2,80	2,15E-06
SORBIDRAFT_02g002260	hypothetical protein		2,75	2,37E-07
SORBIDRAFT_04g002643	hypothetical protein		2,73	9,42E-07
SORBIDRAFT_01g034560	hypothetical protein	Domain associated at C-terminal with AAA	2,67	3,43E-07
SORBIDRAFT_06g001610	hypothetical protein	cysteine synthase/L-3-	2,58	4,02E-06

Continuation

		cyanoalanine synthase		
SORBIDRAFT_01g009930	hypothetical protein		2,50	2,95E-08
SORBIDRAFT_04g002640	hypothetical protein		2,50	2,21E-12
SORBIDRAFT_05g002720	hypothetical protein	UBL5 ubiquitin-like modifier	2,50	7,98E-07
SORBIDRAFT_03g043430	hypothetical protein		2,41	7,73E-09
		Down Regulated		
SORBIDRAFT_07g004110	hypothetical protein	Glycosyltransferase family A	-309,17	1,33E-16
SORBIDRAFT_02g040060	hypothetical protein	Putative thiazole synthesi	-222,74	2,81E-71
SORBIDRAFT_01g030440	hypothetical protein	N-terminal C2 in EEIG1 and EHBP1 proteins	-208,21	2,29E-11
SORBIDRAFT_04g006270	hypothetical protein	Peroxiredoxin (PRX) family	-150,87	1,49E-12
SORBIDRAFT_01g034930	hypothetical protein	Cytochrome P450	-149,78	9,75E-07
SORBIDRAFT_07g024290	hypothetical protein	fatty acyl CoA reductases (FARs)	-139,64	5,22E-11
SORBIDRAFT_02g041150	hypothetical protein		-96,56	4,01E-07
SORBIDRAFT_06g026640	hypothetical protein		-75,03	1,00E-03
SORBIDRAFT_10g003740	hypothetical protein		-61,54	1,74E-05
SORBIDRAFT_04g003470	hypothetical protein	similar to Putative stress resistance-related protein	-55,30	3,94E-09
SORBIDRAFT_01g049120	hypothetical protein	Transcription factor homologous to NACalpha-BTF3	-53,08	1,93E-07
SORBIDRAFT_02g028370	hypothetical protein	Cupin-like domain	-47,58	7,53E-07
SORBIDRAFT_06g026510	hypothetical protein	Serine/threonine phosphatases	-38,73	8,71E-05
SORBIDRAFT_08g002670	hypothetical protein	nsLTP1	-37,31	2,84E-05
SORBIDRAFT_07g003570	hypothetical protein	ISOPREN_C2_like	-36,83	1,14E-06
SORBIDRAFT_05g026200	hypothetical protein		-32,87	1,35E-04
SORBIDRAFT_02g006290	hypothetical protein	cellulose synthase A (UDP-forming)	-32,71	3,01E-35
SORBIDRAFT_06g026630	hypothetical protein	zinc-finger of the FCS-type	-31,87	2,10E-08
SORBIDRAFT_06g032790	hypothetical protein	HSP70 interaction site	-31,10	2,98E-16
SORBIDRAFT_08g019300	hypothetical protein	ISOPREN_C2_like	-28,89	1,41E-07
SORBIDRAFT_01g020830	hypothetical protein		-28,55	4,22E-04
SORBIDRAFT_02g033590	hypothetical protein	Ferredoxin	-27,65	1,70E-19
SORBIDRAFT_03g037490	hypothetical protein	tubulin beta chain	-26,31	1,42E-04
SORBIDRAFT_01g010600	hypothetical protein	FKBP-type peptidyl-prolyl cis-trans isomerase	-26,13	1,25E-12
SORBIDRAFT_02g030990	hypothetical protein	endoglucanase	-23,63	1,45E-09
SORBIDRAFT_04g026190	hypothetical protein	CheY chemotaxis protein	-22,78	1,02E-04
SORBIDRAFT_09g026130	hypothetical protein		-22,19	9,21E-11
SORBIDRAFT_10g001600	hypothetical protein	Ubiquitin-like domain of BAG1	-21,82	1,25E-10
SORBIDRAFT_01g049750	hypothetical protein	Protein Kinases	-20,71	2,18E-10
SORBIDRAFT_01g012050	hypothetical protein	ThiC-associated domain	-20,06	6,14E-19
SORBIDRAFT_03g003310	hypothetical protein	DDE superfamily endonuclease	-18,68	1,02E-12
SORBIDRAFT_05g008570	hypothetical protein	NBD_sugar-kinase_HSP70_actin	-17,80	4,58E-06
SORBIDRAFT_09g005280	hypothetical protein	cellulose synthase	-17,76	1,08E-34
SORBIDRAFT_01g049810	hypothetical protein	Remorin, C-terminal region	-17,27	1,23E-06

Continuation

SORBIDRAFT_09g028490	hypothetical protein	Fasciclin domain	-17,21	9,12E-18
SORBIDRAFT_10g004800	hypothetical protein	Wall-associated receptor kinase galacturonan-binding	-16,80	4,54E-06
SORBIDRAFT_01g008570	hypothetical protein	Helix-loop-helix domain	-16,68	2,17E-07
SORBIDRAFT_03g047330	hypothetical protein	myb-like DNA-binding domain	-16,16	1,97E-06
SORBIDRAFT_08g001780	hypothetical protein	Phytocyanins	-16,12	1,69E-04
SORBIDRAFT_05g026100	hypothetical protein	"Sodium/calcium exchanger protein	-15,48	2,82E-07
SORBIDRAFT_02g006680	hypothetical protein		-14,31	1,11E-11
SORBIDRAFT_01g044010	hypothetical protein	Sugar (and other) transporter	-13,09	1,97E-07
SORBIDRAFT_01g025570	hypothetical protein	inc-dependent alcohol dehydrogenase-like family	-12,67	1,65E-06
SORBIDRAFT_01g005770	hypothetical protein	Fasciclin domain	-11,79	1,95E-06
SORBIDRAFT_10g006380	hypothetical protein		-11,34	1,02E-06
SORBIDRAFT_01g015270	hypothetical protein		-11,11	5,52E-11
SORBIDRAFT_01g042490	hypothetical protein	Caleosin related protein	-11,07	8,25E-05
SORBIDRAFT_09g014430	hypothetical protein	Heat shock protein 60	-11,06	4,12E-05
SORBIDRAFT_04g007970	hypothetical protein	Pimeloyl-ACP methyl ester carboxylesterase	-11,06	9,70E-06
SORBIDRAFT_04g009090	hypothetical protein		-10,63	9,59E-07
SORBIDRAFT_03g040060	hypothetical protein	GT1_Glycogen_Phosphorylase	-10,42	1,32E-28
SORBIDRAFT_04g027550	hypothetical protein		-10,21	5,71E-06
SORBIDRAFT_06g016555	hypothetical protein		-9,89	3,34E-04
SORBIDRAFT_09g016840	hypothetical protein	chlorophyll synthetase	-9,85	1,51E-03
SORBIDRAFT_03g007910	hypothetical protein		-9,80	2,02E-05
SORBIDRAFT_03g025190	hypothetical protein	MFS	-9,76	1,07E-21
SORBIDRAFT_01g009570	hypothetical protein	tubulin alpha chain	-9,29	2,38E-05
SORBIDRAFT_01g036240	hypothetical protein	photosystem II protein Psb27	-9,21	3,41E-05
SORBIDRAFT_06g023880	hypothetical protein	Kinesin motor domain, CENP-E/KIP2-like subgroup	-8,74	2,50E-09
SORBIDRAFT_04g029940	hypothetical protein		-8,64	1,83E-05
SORBIDRAFT_03g039050	hypothetical protein		-8,33	7,91E-06
SORBIDRAFT_07g018840	hypothetical protein	UDP-glucose 4-epimerase	-8,31	1,69E-06
SORBIDRAFT_03g037380	hypothetical protein		-8,23	5,81E-08
SORBIDRAFT_06g032780	hypothetical protein	NUMOD3 motif (2 copies)	-8,08	3,51E-05
SORBIDRAFT_10g006700	hypothetical protein	Myb-like DNA-binding domain	-8,00	2,48E-06
SORBIDRAFT_10g026420	hypothetical protein	crypto_DASH	-7,94	4,43E-04
SORBIDRAFT_08g023070	hypothetical protein	Cation_ATPase_N	-7,86	5,59E-05
SORBIDRAFT_07g006195	hypothetical protein	DNA-binding domain	-7,86	4,21E-04
SORBIDRAFT_09g029610	hypothetical protein	glucose-1-phosphate adenylyltransferase	-7,85	2,44E-12
SORBIDRAFT_01g003710	hypothetical protein	No apical meristem (NAM) protein	-7,84	1,69E-06
SORBIDRAFT_10g007840	hypothetical protein	C2 domain	-7,81	1,74E-05
SORBIDRAFT_06g030240	hypothetical protein		-7,81	2,01E-04

Continuation

SORBIDRAFT_02g003900	hypothetical protein	RNase_PH	-7,73	1,21E-04
SORBIDRAFT_03g008240	hypothetical protein		-7,71	1,84E-07
SORBIDRAFT_01g007760	hypothetical protein	COBRA-like protein	-7,68	1,12E-17
SORBIDRAFT_09g020140	hypothetical protein	1-deoxy-D-xylulose-5-phosphate synthase	-7,37	4,15E-05
SORBIDRAFT_03g038330	hypothetical protein		-7,31	3,66E-09
SORBIDRAFT_09g002740	hypothetical protein	secretory_peroxidase	-7,26	2,86E-07
SORBIDRAFT_06g015360	hypothetical protein		-7,24	1,16E-36
SORBIDRAFT_02g019490	hypothetical protein	phosphoglucosamine mutase family protein	-7,22	1,39E-03
SORBIDRAFT_01g002610	hypothetical protein	50S ribosomal protein L18	-7,10	1,07E-03
SORBIDRAFT_02g007820	hypothetical protein		-7,09	1,59E-04
SORBIDRAFT_03g040630	hypothetical protein	Glycosyl hydrolase family 1	-7,06	4,41E-05
SORBIDRAFT_10g022680	hypothetical protein	Cyclopropane fatty-acyl-phospholipid synthase	-7,03	5,16E-05
SORBIDRAFT_08g005410	hypothetical protein	GDP-L-galactose-hexose-1-phosphate guanyltransferase	-7,01	4,73E-12
SORBIDRAFT_01g012740	hypothetical protein	tubulin beta chain	-6,98	3,74E-07
SORBIDRAFT_10g007850	hypothetical protein		-6,97	1,26E-18
SORBIDRAFT_01g018230	hypothetical protein	short chain dehydrogenase	-6,72	1,06E-07
SORBIDRAFT_02g034980	hypothetical protein	Tpt phosphate/phosphoenolpyruvate translocator	-6,68	5,15E-10
SORBIDRAFT_04g030360	hypothetical protein		-6,66	1,54E-07
SORBIDRAFT_03g047270	hypothetical protein	Drought induced 19 protein (Di19)	-6,62	1,98E-04
SORBIDRAFT_10g001120	hypothetical protein	Chaperonin-60 beta subunit	-6,54	2,04E-13
SORBIDRAFT_03g041100	hypothetical protein	photosystem II subunit S	-6,50	9,78E-28
SORBIDRAFT_01g045010	hypothetical protein	Leucine rich repeat N-terminal domain	-6,46	1,10E-03
SORBIDRAFT_08g002660	hypothetical protein	nsLTP1	-6,43	5,10E-09
SORBIDRAFT_03g045420	hypothetical protein	hexokinase	-6,31	1,23E-04
SORBIDRAFT_02g024760	hypothetical protein	Rab interacting lysosomal protein-like 1 and 2	-6,22	6,72E-06
SORBIDRAFT_03g045440	hypothetical protein	Mpp10 protein	-5,98	1,18E-03
SORBIDRAFT_02g031460	hypothetical protein	pepsin_retropepsin_like	-5,98	6,15E-04
SORBIDRAFT_01g008860	hypothetical protein	Glycosyl hydrolase family 9	-5,96	7,76E-11
SORBIDRAFT_01g041150	hypothetical protein	Filament-like plant protein, long coiled-coil	-5,88	1,85E-09
SORBIDRAFT_02g037970	hypothetical protein	Formin Homology 2 Domain	-5,87	5,28E-07
SORBIDRAFT_02g041940	hypothetical protein	Ribosomal protein S1-like RNA-binding domain	-5,86	3,21E-15
SORBIDRAFT_04g036380	hypothetical protein	50S ribosomal protein L9	-5,85	2,12E-04
SORBIDRAFT_08g003300	hypothetical protein		-5,76	1,49E-03
SORBIDRAFT_06g017130	hypothetical protein	Regulator of chromosome	-5,71	1,29E-05

Continuation

		condensation (RCC1) repeat		
SORBIDRAFT_02g040260	hypothetical protein	RNA recognition motif (RRM) domain	-5,64	4,10E-19
SORBIDRAFT_07g026620	hypothetical protein	60S ribosomal protein L32;	-5,63	8,39E-04
SORBIDRAFT_10g009560	hypothetical protein	CAAD domains of cyanobacterial aminoacyl-tRN	-5,56	3,08E-15
SORBIDRAFT_08g005390	hypothetical protein		-5,49	1,15E-06
SORBIDRAFT_09g028690	hypothetical protein	Catalytic domain of the Serine/Threonine Kinases	-5,49	4,75E-15
SORBIDRAFT_06g032110	hypothetical protein	Heavy-metal-associated domain	-5,40	3,18E-06
SORBIDRAFT_06g018330	hypothetical protein	alpha/beta hydrolases	-5,33	2,38E-04
SORBIDRAFT_10g007330	hypothetical protein	Old yellow enzyme (OYE)-like FMN binding domain	-5,30	2,41E-09
SORBIDRAFT_10g023850	hypothetical protein	fructose-bisphosphate aldolase	-5,21	4,38E-11
SORBIDRAFT_07g027880	hypothetical protein		-5,17	1,28E-06
SORBIDRAFT_10g025480	hypothetical protein	Ribosomal_L35p	-5,11	1,45E-04
SORBIDRAFT_05g001760	hypothetical protein	Na+-driven multidrug efflux pump [Defense mechanisms	-5,04	5,14E-14
SORBIDRAFT_02g003510	hypothetical protein	Phospholipid methyltransferase	-5,04	3,56E-08
SORBIDRAFT_01g015580	hypothetical protein	Starch/carbohydrate-binding module (family 53)	-5,01	7,39E-13
SORBIDRAFT_07g006090	hypothetical protein	probable cinnamyl alcohol dehydrogenase	-4,95	1,32E-04
SORBIDRAFT_03g035290	hypothetical protein		-4,91	1,13E-04
SORBIDRAFT_04g034660	hypothetical protein		-4,87	1,55E-04
SORBIDRAFT_04g037070	hypothetical protein	PDZ domain of C-terminal processing	-4,87	5,44E-04
SORBIDRAFT_02g005440	hypothetical protein	Cation_ATPase_N	-4,82	2,33E-04
SORBIDRAFT_04g028550	hypothetical protein	His Kinase A	-4,80	1,11E-04
SORBIDRAFT_04g028460	hypothetical protein	Glycosyl transferase family 8	-4,71	3,09E-10
SORBIDRAFT_04g036413	hypothetical protein	CLP_protease	-4,61	3,16E-05
SORBIDRAFT_03g010630	hypothetical protein	DNA gyrase subunit B	-4,61	1,42E-03
SORBIDRAFT_05g023520	hypothetical protein	alt stress response/antifungal	-4,58	5,19E-06
SORBIDRAFT_10g006530	hypothetical protein		-4,55	3,03E-06
SORBIDRAFT_01g002050	hypothetical protein	Zinc-binding RING-finger	-4,49	1,98E-10
SORBIDRAFT_06g022540	hypothetical protein		-4,37	1,14E-04
SORBIDRAFT_01g007340	hypothetical protein	Serine/threonine phosphatases	-4,34	9,82E-06
SORBIDRAFT_01g007050	hypothetical protein	PetM family of cytochrome b6f complex subunit 7	-4,27	1,49E-04
SORBIDRAFT_03g026510	hypothetical protein	light-harvesting-like protein 3	-4,23	6,54E-05
SORBIDRAFT_01g037280	hypothetical protein	EF-hand domain pair	-4,20	1,54E-04
SORBIDRAFT_02g023170	hypothetical protein	zinc-finger of the FCS-type	-4,17	4,66E-08
SORBIDRAFT_01g011305	hypothetical protein	chloroplast Hsp70	-4,07	9,85E-04
SORBIDRAFT_07g022140	hypothetical protein	zinc-finger of the FCS-type	-4,06	2,62E-09
SORBIDRAFT_05g019766	hypothetical protein		-3,93	3,64E-04

Conclusion

SORBIDRAFT_02g039210	hypothetical protein	similar to COP1-interacting protein 7 (CIP7)-like	-3,93	3,93E-04
SORBIDRAFT_01g042380	hypothetical protein	alpha,alpha-trehalose-phosphate synthase	-3,82	1,43E-04
SORBIDRAFT_01g001990	hypothetical protein	Ligand-binding SRPBCC domain of an uncharacterized	-3,64	6,43E-05
SORBIDRAFT_09g029900	hypothetical protein	ATPases	-3,54	2,14E-04
SORBIDRAFT_07g014860	hypothetical protein		-3,50	1,02E-07
SORBIDRAFT_02g031930	hypothetical protein		-3,48	4,70E-05
SORBIDRAFT_07g021760	hypothetical protein	Omega-hydroxypalmitate O-feruloyl transferase	-3,44	3,25E-08
SORBIDRAFT_01g018450	hypothetical protein	Chloroplast envelope transporter	-3,42	6,65E-05
SORBIDRAFT_06g030160	hypothetical protein	glutamyl-tRNA reductase	-3,34	4,05E-04
SORBIDRAFT_02g024550	hypothetical protein	EF-hand, calcium binding motif	-3,27	4,77E-08
SORBIDRAFT_02g039660	hypothetical protein	phospho-2-dehydro-3-deoxyheptonate aldolase	-3,22	1,08E-03
SORBIDRAFT_09g029580	hypothetical protein		-3,21	9,40E-05
SORBIDRAFT_06g023840	hypothetical protein	EF-G	-3,15	1,87E-12
SORBIDRAFT_01g035090	hypothetical protein	50S ribosomal protein L6	-3,13	5,62E-06
SORBIDRAFT_04g005520	hypothetical protein	WRKY DNA binding domain	-3,13	1,80E-07
SORBIDRAFT_03g034370	hypothetical protein	Chromatin remodeling complex protein RSC6	-3,10	1,73E-11
SORBIDRAFT_05g004590	hypothetical protein	TIM_phosphate_binding	-3,04	5,53E-04
SORBIDRAFT_07g019300	hypothetical protein		-3,02	7,38E-04
SORBIDRAFT_10g007180	hypothetical protein	CP12 domain	-3,00	1,26E-11
SORBIDRAFT_10g017820	hypothetical protein	Golgi-localized syntaxin-1-binding clamp	-2,97	6,99E-07
SORBIDRAFT_04g011160	hypothetical protein	TypA_BipA	-2,96	1,60E-07
SORBIDRAFT_02g028570	hypothetical protein	SPA9-like_NBD	-2,91	7,68E-06
SORBIDRAFT_01g039530	hypothetical protein	heat shock 70 kDa protein	-2,85	4,77E-07
SORBIDRAFT_07g025000	hypothetical protein	Ribosomal protein S19e	-2,84	3,64E-05
SORBIDRAFT_03g007000	hypothetical protein	Calmodulin binding protein-like	-2,83	5,44E-05
SORBIDRAFT_03g000370	hypothetical protein	Tpt phosphate/ phospho enolpyruvate translocator;	-2,79	1,14E-07
SORBIDRAFT_03g010620	hypothetical protein	MFS	-2,71	1,85E-10
SORBIDRAFT_01g050480	hypothetical protein	1,3-beta-glucan synthase subunit FKS1	-2,70	4,06E-12
SORBIDRAFT_06g001130	hypothetical protein	Lipid-binding START domain of mammalian STARD2	-2,70	2,13E-04
SORBIDRAFT_01g021890	hypothetical protein	GDP-D-mannose-3',5'-epimerase	-2,63	2,21E-06
SORBIDRAFT_07g028270	hypothetical protein	heat shock protein 83 kDa (Hsp83	-2,58	1,68E-05
SORBIDRAFT_09g019390	hypothetical protein	S-adenosylmethionine-dependent methyltransferases	-2,56	5,24E-11
SORBIDRAFT_08g004500	hypothetical protein	TIM_phosphate_binding	-2,49	9,64E-12

SOURCE: The Author (2017).

CAPÍTULO V

Manuscrito ainda não submetido

New insights in Plant Bacteria Interaction: Study of the beneficial association between *Herbaspirillum seropedicae* and *Setaria Viridis* using RB-TnSeq analysis

New insights in Plant Bacteria Interaction: Study of the beneficial association between *Herbaspirillum seropedicae* and *Setaria Viridis* using RB-TnSeq analysis

ABSTRACT

Herbaspirillum seropedicae is a nitrogen fixing bacteria that can be found in association with several important crops, as maize, rice and wheat. This association can increase significantly the nitrogen available for the plants and promote the plant growth. Besides the Nitrogen, this beneficial association might involve several others compounds, as hormones or antibiotics to protect the plant against pathogens. Even though this association is very important there are still several processes that remain unknown. In this work we analyzed the association between *H. seropedicae* and a grass model, *Setaria viridis*, using a new approach for plant bacteria interaction studies, the BarSeq. A pool of mutants was inoculated in *Setaria viridis* plants and the fitness of each gene showed us some essential candidate genes for the association between plant and bacteria. Several genes involved with chemotaxis were identify, showing that this process is important to the microbe find the plant and stabilize the interaction. The flagellum assembly genes were also found, but the mutations promote a better colonization of *H.seropedicae* and *Setaria viridis*, suggesting that the plant can have a delay in the microbe recognition and the bacteria could colonize more successfully. Genes involved with PHB metabolism also showed different behaviors and their involvement with plant bacteria interaction can be strictly modulate by the association with plants. Using this technique we could generate a high amount of data and a deeply analysis of all this candidates genes suggested some important pathways related with plant bacteria interaction.

INTRODUCTION

The interaction between plants and diazotrophic bacteria has been more studied in the last few years, as well as the use of endophytic bacteria, such as *Rhizobium, Azospirillum, Gluconacetobacter, Azoarcus,* and *Herbaspirillum* as plant growth promoting rhizobacteria (PGPR) (MONTEIRO et al., 2012). This bacterial group is important since they can colonize different tissues, promote the plant growth by the control of phytopathogens, increase the uptake of minerals and the production of different phytohormones, besides the nitrogen fixation (ELBELTAGY et al., 2001; INIGUEZ et al., 2004; SESSITSCH et al., 2002; STURZ et al., 2000).

The *Herbaspirillum* genus comprises Betaproteobacteria with different life style; some species live in aquatic environments, soil or can be found as opportunistic pathogens in human diseases (BERG et al., 2005). This opportunistic species are not capable of fixing nitrogen. In contrast, some species can fix the atmospheric nitrogen, under microaerobic conditions, such as *Herbaspirillum seropedicae* and *Herbaspirillum rubrisubalbicans*, and can associate beneficially with different crops. *Herbaspirillum seropedicae* is found in association with several economically important crops, such as maize, rice and sugar cane (BALDANI et al., 1986; BALDANI et al., 1992; BODDEY et al., 1995; JAMES et al., 2000; JAMES et al., 2002; ELBELTAGY et al., 2001; GYANESHWAR et al., 2002; RONCATO-MACCARI et al., 2003). The association with *H. seropedicae* is normally beneficial to the plant, increasing the accumulated nitrogen up to 120%, and, consequently, increasing the plant growth (BALDANI et al., 2000).

The association between *H. seropedicae* and plants is still poorly understood, but some important factors were already elucidated, as the importance of polysaccharides biosynthesis, the alteration on the bacterial cell envelope (synthesis of LPS, protein secretion components and alteration on the peptidoglycans portions) and the PHB metabolism during the interaction with maize (BALSANELLI et al., 2015).

The interaction between *H. seropedicae* (co-inoculated with *Azospirillum brasilense*) and *Setaria viridis* showed a high number of bacterial colonization, significantly increase in seed number, shoot fresh/dry weight, root length and number of lateral root (40 days after inoculation) (PANKIEVICZ et al., 2015).

To better understanding the plant bacteria interaction process, one of the most powerful techniques can be used is the TnSeq (Transposon Sequencing). This approach is based on the use of transposons, mobile genetic elements, to obtain random insertion mutagenesis. Combining this technique with the new generation sequencing platforms it is possible to follow the mutagenized regions and their respective genes and functions and thus assume the importance of the genes for the condition in which the bacterium is being tested (OPIJNEN AND CAMILLI, 2013).

Among the different methodologies used to perform the TnSeq is the Bar-Seq. This methodology allows associating random mutagenesis (Tn-Seq) and the use of BarCodes - DNA sequences - previously associated with specific mutations. This previous association works as genetic markers in a single pool of mutants. In this way it is possible to estimate the importance of each gene for a specific condition by counting the number of reads that correspond to the mapped Barcodes (WETMORE et al., 2015).

Several works with microorganisms have already been developed using Tn-Seq technique including studies with pathogenic organisms, such as *Vibrio cholerae* and *Pseudomonas aeruginosa*, where important factors for the pathogenesis were found as effector proteins of the Type VI Secretion System in *V. cholerae* (DONG et al., 2012; WETMORE et al., 2015).

Recent studies have demonstrated the efficiency of the Tn-Seq technique associated with the use of DNA Barcodes, which is an important tool for studying the relationship between genes and their functions in a variety of environments (WETMORE et al., 2015).

RESULTS AND DISCUSSION

To better understand the mechanistic process involving the interaction between *H. seropedicare* SmR1 and *Setaria viridis* A10.1 was conducted a TnSeq assay during this interaction. The *Herbaspirillum seropedicae* TnSeq library was constructed as previous described by Wetmore, 2015. The plants were grown for 5 days then we inoculated the mutant pool library with 10 mL of a solution containing 10^8 bacteria per mL. Ten days after the inoculation we recovered the epiphytic population of these plants and extracted their gDNA to sequencing analysis. As control of this experiment we used a pot just with soil. The soil was inoculated in the same way as the *Setaria viridis* plants. We watered all the plants with a Hoaglands Solution (low Nitrogen). As the soil does not have a Carbon source, we watered these pots with a Hoaglands Solution containing 0.1% of Malic Acid each 3 days. The yield of the DNA extraction is shown in table 5-1.

TABLE 5-1 – Yield of total gDNA extractions after the growth in the determined conditions

Samples	DNA (ng/uL)	Total DNA (ug)
SmR1 A10.1 S1	44.81	**8.96**
SmR1 A10.1 S2	39.55	**7.91**
SmR1 A10.1 S3	53.72	**10.74**
SmR1 A10.1 S4	71.74	**14.34**
SmR1 A10.1 S5	102.70	**20.54**
SmR1 SOIL 1	37.20	**7.44**
SmR1 SOIL 2	37.51	**7.50**

SOURCE: The Author (2017).

The samples were sequenced, mapped according to the BarCode and the gene fitness analysis was made using the package FEBA.R, R program.

According to the data obtained, were found 196 genes with mutations responsible for impairing the epiphytic colonization of *Herbaspirillum seropedicae* SmR1 (SICK GENES - gene fitness ≤ -1), of these 80 genes were also present in the Control (Soil without plant), being excluded of the analysis (Table 5-2).

TABLE 5-2 – Number of SICK genes found during the TnSeq analysis.

SICK GENES		
	PLANT	**SOIL**
Total of SICK genes	196	118
	TOTAL OF GENES	
BOTH (SOIL and PLANT)		80
Exclusively SOIL		38
Exclusively PLANT		116

SOURCE: The Author (2017).

Among the Sick genes, they can be categorized in 3 groups: Incapable of Growth (gene fitness around -6), Very Sick (gene fitness between -2 to -1) and Mildly deleterious (gene fitness ≤ -1). The table 5-3 shows all the genes found and their respectively gene fitness.

TABLE 5-3 – SICK – Genes in which the mutation is responsible for impair the capacity of epiphytic colonization.

Locus Id	Gene Annotation	*Gene fitness*
HSERO_RS18990	pyrroline-5-carboxylate reductase	-6.56398
HSERO_RS07755	glutamate racemase	-5.37653
HSERO_RS03085	3-methyl-2-oxobutanoate hydroxymethyltransferase	-3.75144
HSERO_RS20315	imidazole glycerol phosphate synthase	-3.2727
HSERO_RS07970	formyltetrahydrofolate deformylase	-3.06146
HSERO_RS09220	hypothetical protein	-3.02952
HSERO_RS01265	acetyl-CoA acetyltransferase	-2.92923
HSERO_RS13965	single-stranded DNA exonuclease	-2.79179
HSERO_RS17980	GTP-binding protein	-2.78211
HSERO_RS13960	hypothetical protein	-2.76975
HSERO_RS03445	biopolymer transport transmembrane protein	-2.61733
HSERO_RS09145	DNA topoisomerase IV subunit A	-2.51161
HSERO_RS03440	biopolymer transporter ExbD	-2.4366
HSERO_RS17105	DNA polymerase III subunit epsilon	-2.4279
HSERO_RS02650	dihydroxy-acid dehydratase	-2.39864
HSERO_RS05125	DeoR family transcripitonal regulator	-2.30441
HSERO_RS01880	N-acetyl-anhydromuranmyl-L-alanine amidase	-2.23744
HSERO_RS18600	transcriptional regulator	-2.16223
HSERO_RS10395	phosphate transport regulator	-2.14995
HSERO_RS03520	aminoglycoside phosphotransferase	-2.12597
HSERO_RS14950	purine-binding chemotaxis protein	-2.11656
HSERO_RS23945	cobyric acid synthase CobQ	-2.11238
HSERO_RS09155	DNA topoisomerase IV subunit B	-2.06511
HSERO_RS07025	ATP-dependent Clp protease ClpS	-2.05123
HSERO_RS05540	cytochrome C	-2.03414
HSERO_RS10850	ABC transporter	-2.03307
HSERO_RS14955	chemotaxis protein	-1.99165
HSERO_RS00080	histidine kinase	-1.9458
HSERO_RS20750	cytochrome C oxidase subunit I	-1.94285
HSERO_RS13260	cobinamide adenolsyltransferase	-1.92189
HSERO_RS00430	glutamate--cysteine ligase	-1.89951
HSERO_RS13305	magnesium chelatase	-1.89052
HSERO_RS09235	MarR family transcriptional regulator	-1.86426
HSERO_RS18010	pseudouridine synthase	-1.80458
HSERO_RS13265	G3E family GTPase	-1.77664
HSERO_RS01870	ABC transporter permease	-1.75741
HSERO_RS10125	chemotaxis protein	-1.75693
HSERO_RS19285	long-chain fatty acid--CoA ligase	-1.74599
HSERO_RS08535	glycine/betaine ABC transporter substrate-binding protein	-1.71237
HSERO_RS14960	Fis family transcriptional regulator	-1.67789
HSERO_RS16505	3-isopropylmalate dehydratase small subunit	-1.6671

Continuation

HSERO_RS10115	chemotaxis protein CheR	-1.62865
HSERO_RS03270	protein RecA	-1.62138
HSERO_RS07290	oligopeptidase A	-1.61386
HSERO_RS19450	indole-3-glycerol-phosphate synthase	-1.60792
HSERO_RS13315	ferredoxin	-1.60435
HSERO_RS08530	glycine/betaine ABC transporter permease	-1.5724
HSERO_RS19135	exodeoxyribonuclease	-1.54035
HSERO_RS13660	membrane protein	-1.53846
HSERO_RS00795	glycerol-3-phosphate dehydrogenase	-1.52556
HSERO_RS12325	arabinose ABC transporter permease	-1.51258
HSERO_RS09050	2-aminoadipate aminotransferase	-1.49275
HSERO_RS14835	peptidase M23	-1.49183
HSERO_RS13255	cobyrinic acid a,c-diamide synthase	-1.49083
HSERO_RS00220	hypothetical protein	-1.48947
HSERO_RS21740	methionine synthase	-1.4687
HSERO_RS19835	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase	-1.45754
HSERO_RS15655	methylcitrate synthase	-1.45237
HSERO_RS20790	cytochrome B559 subunit alpha	-1.4509
HSERO_RS17720	aminopeptidase N	-1.44578
HSERO_RS11610	NADH dehydrogenase	-1.42331
HSERO_RS01825	RNA pyrophosphohydrolase	-1.41461
HSERO_RS01965	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl- meso-diaminopimelate ligase	-1.41383
HSERO_RS16515	isopropylmalate isomerase	-1.40961
HSERO_RS10340	flagellar biosynthesis protein FliS	-1.39798
HSERO_RS11615	hypothetical protein	-1.39603
HSERO_RS13320	membrane protein	-1.38809
HSERO_RS20235	cytochrome C	-1.38319
HSERO_RS08150	phasin family protein	-1.37971
HSERO_RS02325	uracil phosphoribosyltransferase	-1.37865
HSERO_RS16485	tRNA pseudouridine synthase A	-1.36357
HSERO_RS07300	glutamate--cysteine ligase	-1.35441
HSERO_RS20780	cytochrome C oxidase assembly protein	-1.33999
HSERO_RS00455	homoserine O-acetyltransferase	-1.33257
HSERO_RS02205	phosphatase	-1.32453
HSERO_RS21640	recombinase RdgC	-1.31698
HSERO_RS08660	ketol-acid reductoisomerase	-1.31491
HSERO_RS06165	ATP-dependent DNA helicase RecG	-1.30417
HSERO_RS10320	flagellar brake protein YcgR	-1.29966
HSERO_RS21425	5,10-methylenetetrahydrofolate reductase	-1.29866
HSERO_RS20920	3-dehydroquinate synthase	-1.29024
HSERO_RS00765	cytochrome C	-1.26623
HSERO_RS16480	MarR family transcriptional regulator	-1.26258
HSERO_RS12300	hypothetical protein	-1.2555
HSERO_RS16500	3-isopropylmalate dehydrogenase	-1.25478

Conclusion

HSERO_RS15600	cytosol aminopeptidase	-1.24093
HSERO_RS09660	ribonuclease III	-1.23711
HSERO_RS20785	cytochrome oxidase subunit I	-1.22669
HSERO_RS17185	2-isopropylmalate synthase	-1.21619
HSERO_RS16825	NADP-dependent malic enzyme oxidoreductase	-1.17535
HSERO_RS13285	precorrin-2 C20-methyltransferase	-1.17262
HSERO_RS00425	ammonia channel protein	-1.16245
HSERO_RS09745	chemotaxis protein CheY	-1.15306
HSERO_RS19445	anthranilate phosphoribosyltransferase	-1.15194
HSERO_RS03285	succinyl-CoA synthetase subsunit alpha	-1.1456
HSERO_RS11070	nicotinate-nucleotide pyrophosphorylase	-1.14397
HSERO_RS19090	C4-dicarboxylate ABC transporter	-1.12699
HSERO_RS10650	hypothetical protein	-1.10531
HSERO_RS20355	UDP-N-acetylglucosamine 1-carboxyvinyl transferase	-1.0996
HSERO_RS07155	peptidase C69	-1.08125
HSERO_RS06940	23S rRNA methyltransferase	-1.07884
HSERO_RS01165	dihydroneopterin triphosphate pyrophosphatase	-1.0736
HSERO_RS16835	phosphatidylglycerophosphatase	-1.07297
HSERO_RS11155	peptidyl-prolyl cis-trans isomerase	-1.07287
HSERO_RS00435	glutathione synthetase	-1.06244
HSERO_RS07160	phospho-2-dehydro-3-deoxyheptonate aldolase	-1.06006
HSERO_RS08405	membrane protein	-1.05705
HSERO_RS01510	membrane protein	-1.04991
HSERO_RS07110	16S rRNA methyltransferase	-1.04219
HSERO_RS10495	molybdenum cofactor biosynthesis protein MoaE	-1.0397
HSERO_RS10455	aminotransferase	-1.03871
HSERO_RS18685	hypothetical protein	-1.0316
HSERO_RS20770	MFS transporter	-1.02044
HSERO_RS15555	oxidoreductase	-1.01669
HSERO_RS20970	NDP-N-acetyl-D-mannosaminuronate dehydrogenase	-1.01625
HSERO_RS16075	coproporphyrinogen III oxidase	-1.00329

SOURCE: The Author (2017).

The Figure 5-1 shows the COG (Clusters of Orthologous Groups) Classification of all the proteins found during the TnSeq analysis. The most abundant group is E – Amino acid Metabolism and Transport.

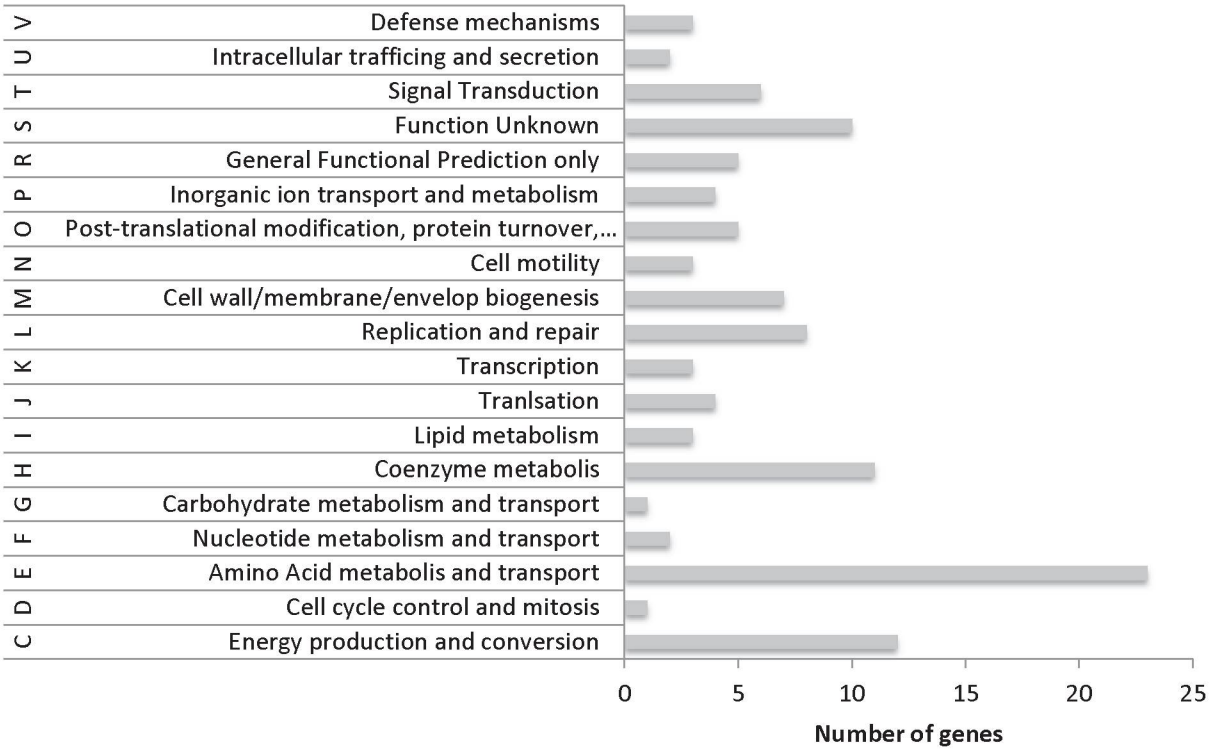


FIGURE 5-1 - COG Classification of the genes found during the TnSeq analysis of *Herbaspirillum seropedicae* colonizing *Setaria viridis*.

Chemotaxis

Several chemotaxis genes were found among the mutations that impair the colonization of *H. seropedicae* in *Setaria viridis: CheR, CheB, CheA, CheW* and *CheY* (FIGURE 5-2). Chemotaxis is directly involved in modulating the movement of the flagellum, helping the bacteria to move to places with more favorable conditions (CLARKE et al., 2016). Many bacteria use chemotaxis pathways to control the motility of the flagellum according to environmental stimulus (BERG and BROWN, 1972). Bacteria oscillate the behavior of "go ahead" (run) and "reorientation". While the forward behavior is favored by the increasing presence of favorable chemicals, the reorientation behavior is favored by the presence of unfavorable chemicals. The result of this balance is the regulation of the flagellum rotation resulting in movements towards the "attractive" and against the "repellents". Chemotaxis has already been described in several organisms as important for epiphytic colonization. Broek (1998) described the importance of the motility and chemotaxis during the interaction between *Azospirillum brasilien*se and wheat.

Herbaspirillum seropedicae has 41 genes involves with chemotaxis (PEDROSA et al., 2011 and TADRA-SFEIR et al., 2015) and some of them were described by Balsanelli (2015) as important to epiphytic and planktonic environments during the association with maize. A mutant in the gene HSERO_RS18565 (Hsero_3720) which encodes a methylaccepting chemotaxis protein (MCP) shows 2 times less bacteria attached to the root, supporting the importance of the chemotaxis to the adaptation of *H. seropedicae* in planta and suggesting that HSERO_RS18565 is an important factor for the recognition of the rhizosphere and for the transduction of the signals from the host-secreted compounds to the bacteria (BALSANELLI et al., 2015). We found 5 different genes related to the chemotaxis process (FIGURE 5-2). The protein CheA (HSERO_RS15535, gene fitness = -1.2) suffer auto phosphorylation induced by MCPs and also induce CheY (HSERO_RS09745, gene fitness = -1.15) phosphorylation. The proteins CheR (HSERO_RS10115, gene fitness= -1.6), CheB and CheW (HSERO_RS14950, gene fitness = -2.1) acts modulating the phosphorylation state of CheY (LI AND HAZELBAUER, 2004). Mutations on this 5 genes impair the capacity of *H. seropedicae* colonize epiphytocally *Setaria viridis,* this results shows the importance of the chemeotaxis to the bacteria recognition of plant environment. As demonstrated by Balsanelli (2015), and Broeck (2008) and supported by this work, the bacteria uses this chemosystem to recognize and establish an interaction with the host.

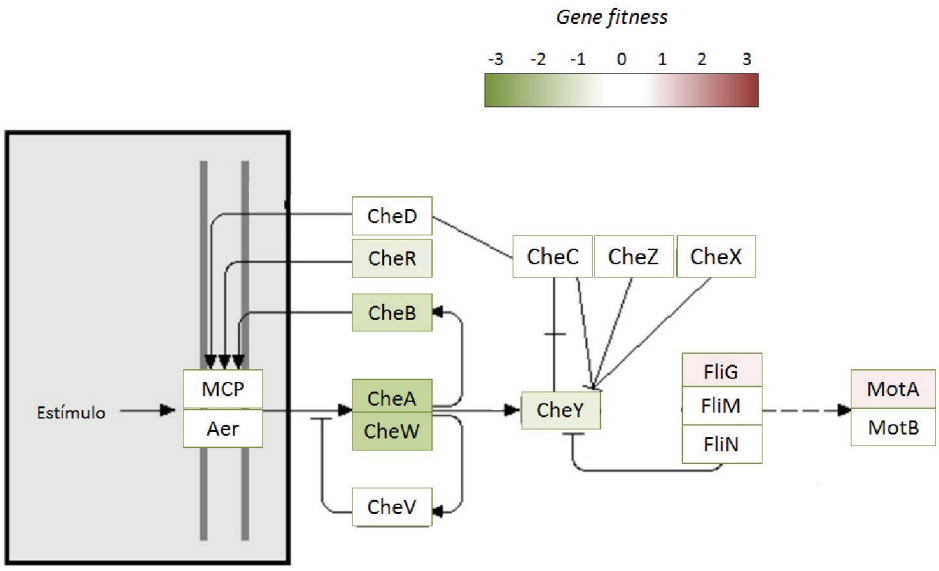


FIGURE 5-2- Bacterial chemotaxis regulation pathway. *che* genes are responsible for transferring the external environmental signals and modulating the activity and direction of the bacterial flagellum rotation. The green proteins are proteins where the mutation in the correspondent gene is responsible for decreasing the fitness gene (the ability of epiphytic colonization), being therefore important proteins for the interaction between *H. seropedicae* and *Setaria viridis*.

Cell Wall

The peptidoglycans are important cell wall components. In *Herbaspirillum seropedicae* the peptidoglycan biosynthesis is encoded by the *mur* genes and their importance in the Plant Bacteria Interaction was reported by Tadra-Sfeir (2015), where several *mur* genes were modulated in the presence of naringenin. Also, Cordeiro (2013) showed that *H. seropedicae* GlmU, an enzyme related to the peptidoglycan synthesis, was modulated in the presence of sugar cane extract. In our results we identified the *murI*, a glutamate racemase (HSERO_RS07755), responsible for the conversion of L-glutamate in D-glutamate and the *murA, an* UDP-N-acetylglucosamine 1-carboxyvinyltransferase (HSERO_RS20355) responsible to converts UDP-N-acetiyglucosamine in UDP-N-acetylglucosamine enolpyruvate, being an important enzyme in the peptidoglycan biosynthesis pathway. The *murI* fitness score was -5.3, indicating that this gene is extremely important for the bacteria growth in the plant bacteria interaction, being the D-Glutamate a precursor required for the peptidoglycan biosynthesis. The lipopolysaccharides (LPS) are also found in the outermambrane of gram-negative bacteria. Balsanelli (2013) demonstrate that lectin-like proteins can mediate the association between maize and *H. seropedicae* by the interaction with the N-acetylglucosamine residues from the bacteria LPS. The mutation of the LPS biosynthesis genes or the addition of exogenous N-acetylglucosamine can impair the attachment of *H. seropedicae* in maize. Our results shows that the peptidoglycans and LPS present in the cell wall are important factors to the interaction of *H. seropedicae* with *Setaria viridis*.

Beneficial mutation

Among the Beneficial mutations - genes where the mutation increases colonization (fitness gene ≥ 1) - 18 genes were found during the interaction with plant, of these 4 were also found in control (soil), totaling 14 unique genes in the plant- Bacteria condition (Table 5-4). The Beneficial mutations exclusively found in plant are showed in the table 5-5.

TABLE 5-4 – Number of BENEFICIAL mutations found during the TnSeq analysis.

BENEFICIAL GENES		
	PLANT	**SOIL**
Total of beneficial	18	10
	TOTAL OF GENES	
BOTH (SOIL and PLANT)		4
Exclusively SOIL		6
Exclusively PLANT		14

SOURCE: The Author (2017).

TABLE 5--5 – Genes where the mutation was beneficial (increasing the colonization capacity of *Herbaspirillum seropedicae* SmR1 in *Setaria viridis* A10.1)

Locus Id	**Gene annotation**	**Gene fitness**
HSERO_RS14975	flagellar motor protein MotA	1.05726
HSERO_RS10140	flagellar biosynthesis protein FlhB	1.066778
HSERO_RS10305	flagellar motor switch protein FliG	1.068837
HSERO_RS06285	methyl-accepting chemotaxis protein	1.078113
HSERO_RS02815	HxlR family transcriptional regulator	1.209123
HSERO_RS10150	flagellar biosynthesis regulator FlhF	1.240156
HSERO_RS10255	flagellar biosynthesis protein FliQ	1.310306
HSERO_RS10310	flagellar M-ring protein FliF	1.33403
HSERO_RS23885	XRE family transcriptional regulator	1.418641
HSERO_RS13885	histidine kinase	1.731006
HSERO_RS14985	transcriptional regulator	1.833253
HSERO_RS13890	LuxR family transcriptional regulator	1.858816
HSERO_RS20835	ABC transporter substrate-binding protein	1.938558
HSERO_RS08080	poly[D(-)-3-hydroxyalkanoate] depolymerase	2.809771

SOURCE: The Author (2017).

Between the 14 exclusively beneficial mutations 6 are involved with flagellum biosynthesis: FliF (HSERO_RS10310, gene fitness=1.3), FliG (HSERO_RS10305, gene fitness=1.06), FlhB (HSERO_RS10140, gene fitness=1.06), FliQ (HSERO_RS10255, gene fitness=1.3), MotA (HSERO_RS14975, gene fitness=1.05) and FlhC (HSERO_RS14985, gene fitness=1.8) (FIGURE 5-3 and 5-4). The genes *fliF, fliG, flhB* and *fliQ* are components of the basal body of the flagellum, being part of their structure. *motA* is complex with *motB* and functions as a proton channel to torque generation (BERG et al., 2003). The FlhC protein is a master transcriptional regulatory factor which works in complex with FlhD and modulate

the transcription of several flagellar and non-flagellar genes (LIU et al., 1994). The flagellum was already described as important to the movement of the bacteria during the interaction with plants, mainly in planktonic bacteria, when the bacteria have to move towards the plant. Balsanelli (2015) showed that a mutant in a flagellum gene (*fliC*) of *H. seropedicae* is incapable of initiate the attachment when the inoculation is made without any agitation (the bacteria cannot move towards the plant without the flagellum). When the bacteria attached the plant in 50 rpm, the bacteria is able again to initiate the adhesion process in the root, even though in a smaller number. This demonstrates that the flagellum is important in this first step of interaction, the signaling process that attracts the bacteria to the plant. The *Herbaspirillum seropedicae* flagella operon is also down regulated after the bacteria is already established in the wheat roots, as demonstrated by Pankievicz (2016). Balsanelli (2015) also demonstrated that the number of epiphytic and entophytic bacteria is the same in the *fliC* mutant and in the wild type after the first attachment moment. Our results shows that the absence of the flagellum increases the ability of the epiphytic colonization in *Setaria viridis* 10 days after inoculation. This suggest that the flagellum is not important once the bacteria is already attached in the root and that the flagella absence can confer an advantage to the bacteria in this stage. Our suggestion is that this advantage can occur since the plants, including sorghum, have specific receptors to sense the bacterial flagellum and start a defense response controlling the number of bacteria attached. If the bacteria do not have the flagellum, probably this defense apparatus is not activated and the bacteria can colonize more succefully than the wild type even after the first interaction step is impaired.

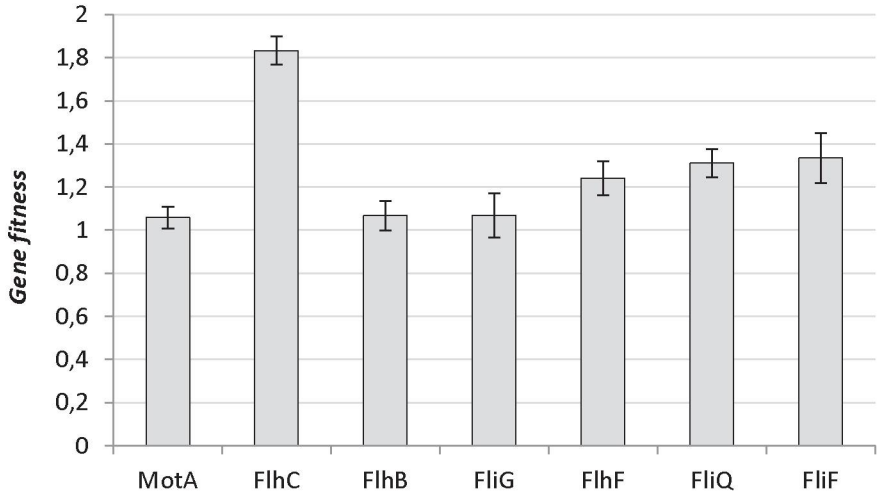


FIGURE 5-3 – *Gene fitness* of genes involved in the flagellum assembly during the assay with *Herbaspirillum seropedicae* SmR1 epiphytically colonizing Setaria viridis A10.1.

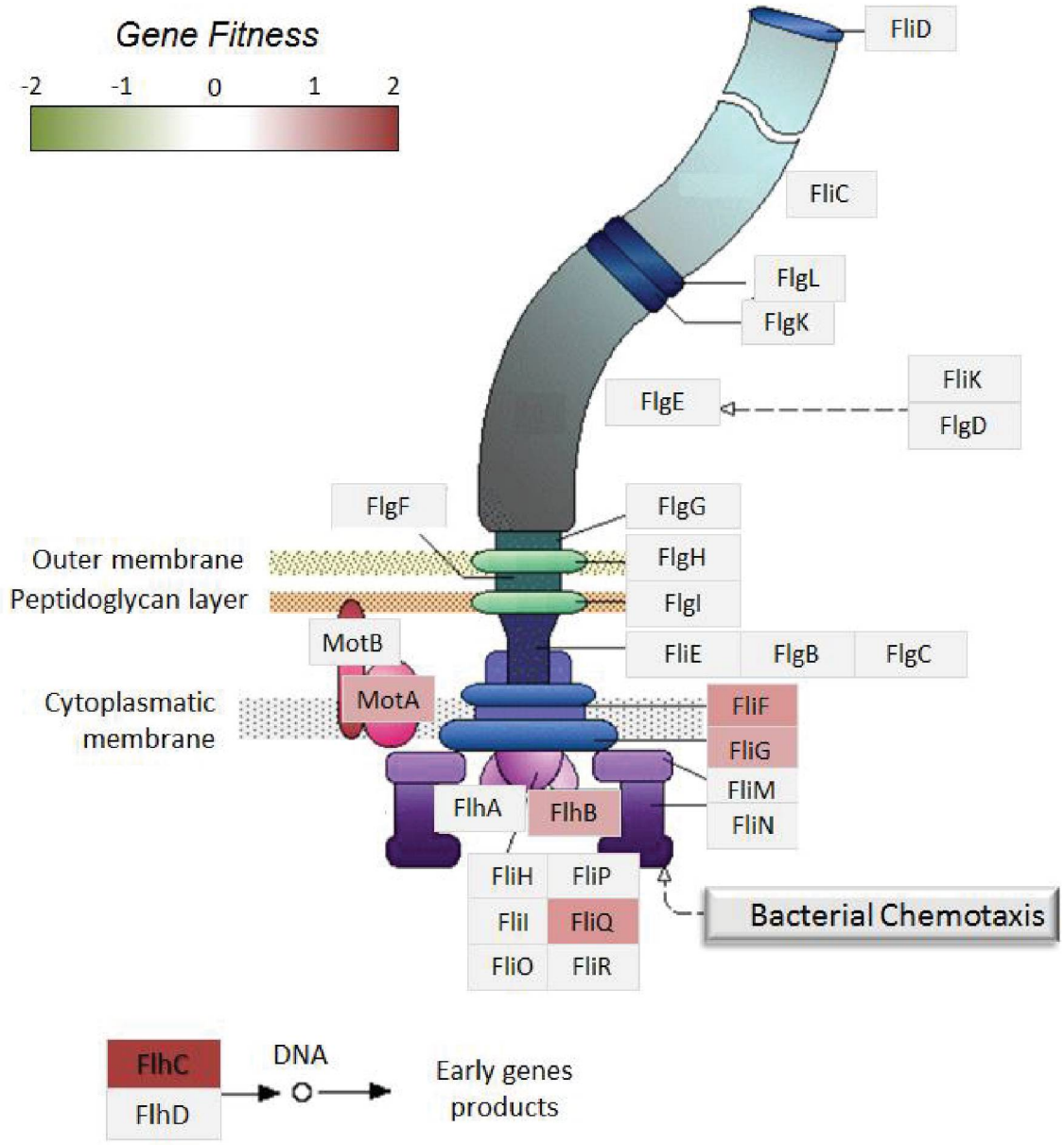


FIGURE 5-4 – Flagellum assembly. FliQ, FlhB, FliG and FliF compound the basal body of the flagellum. MotA is a complex with MotB has function as a proton channel to the flagella movement, FlhC is an important transcriptional regulator of all the flagella assembly. All this genes were found in the TnSeq analysis, being their mutations beneficial to the bacterial epiphytic colonization. Adapted from KEGG.

PHB metabolism

Another 3 important proteins found during this analysis were the PhaZ1 poly-3-hydroxyalkanoate depolymerase (HSERO_RS08080, gene fitness = 2.8), PhaP1 (phasin family protein, HSERO_RS08150, gene fitness = -1.4) and PhbA1 (acetyl-CoA acetyltransferase, HSERO_RS01265, gene fitness = -2.9). The PHB metabolism is responsible for the storage of insoluble intracellular granules of Poly-3-hydroxybutyrate (PHB). This granule acts as storage of Carbon source which can be mobilized in different conditions. *H. seropedicae* has 13 genes involved in the PHB metabolism (KADOWAKI et al., 2011). Between these genes are *phaC* (PHB synthase) required for the synthesis of PHB, *phbA1* and *phbA2* which are alternative biosynthesis genes and *phaZ*, which is a PHB depolymerizes required for the PHB mobilization (BABEL et al., 2001). The phasins (PhaP) are small amphiphilic proteins presents on the border of the granules and have the function of control the size and number of the PHB granule (NEUMANN et al., 2008; JURASEK et al., 2002). The biosynthesis normally occurs in high carbon availability and low presence of essential nutrients, as nitrogen and oxygen (HERVAS et al., 2008). Balsanelli (2015) showed that the PHB metabolism is important during the early stages of *H. seropedicae* colonization in maize, becoming less important in advanced times of interaction between plant and bacteria. Pankievicz (2015) also showed that the phasin proteins and PHB synthases were upregulated during the interaction of *H. seriopedicae* with wheat, showing the importance of the PHB metabolism during the interaction. In our TnSeq results, we found 3 genes related to PHB metabolism in which the mutation was important for the strain fitness: PhaP1 (HSERO_RS08150), PhbA1 (HSERO_RS01265) and PhaZ1 (HSERO_RS08080) (FIGURE 5-5). While PhaP1 data support the results already showed by Balsanelli (2015) and Pankieviks (2016), being this mutant strains less capable of colonize the *Setaria* roots, showing their importance during the interaction; the mutant PhaZ1 had their gene fitness increase when compared with the control (FIGURE 5-5). To confirm the data, we did an Epiphytic colonization assay with the PhaP1 (TIRAPELLE et al., 2013) and PhaZ1 mutants in competition with SmR1-dsRED in the proportion of 1:1. The assay shows a decrease in the number of epiphytic bacteria in the mutant PhaP1 compared with SmR1-dsRED in the competition of all the 3 days analyzed (attachment, 1 day after inoculation and 5 days after inoculation); the decrease ranged between 2.3 and 3.3 times less *ΔPhaP1*. Meanwhile, for the experiment with *ΔPhaZ1*, the mutant increased the number of bacteria during the competition,

the increase ranged between 4.3 and 19.3 times more *ΔPhaZ1* than SmR1-dsRED (Figure 5-6).

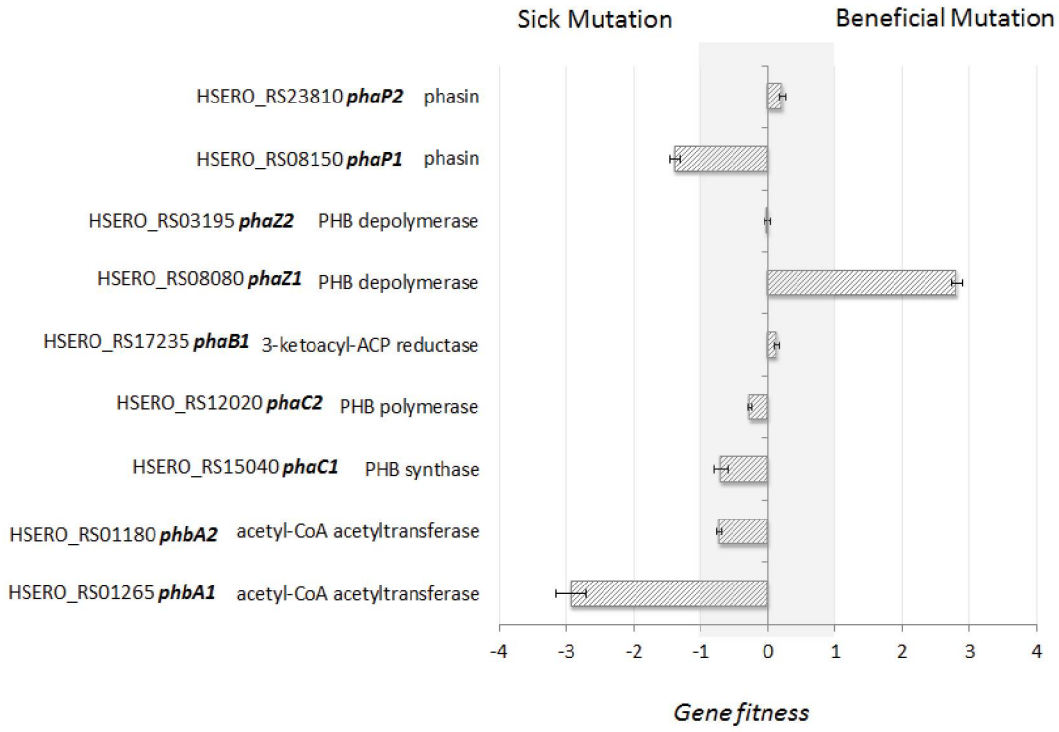


FIGURE 5-5 – Gene fitness of PHB metabolism genes from *H. seropedicae* during the epiphytic colonization in Setaria viridis.

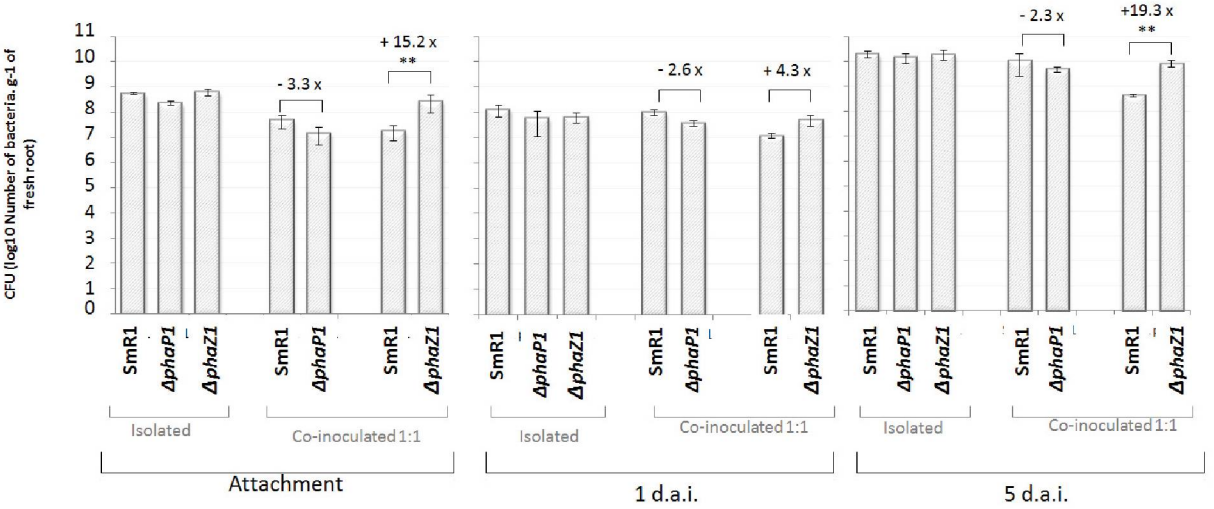


FIGURE 5-6 – Number of epiphytic bacteria colonizing *Setaria viridis* tissue. The assay was performed as attachment (30 minutes after inoculation), 1 d.a.i. (1 day after inoculation) and 5 d.a.i (5 days after inoculation). The strains were inoculated isolated and co-inoculated in the proportion of 1:1. ** Statistically different with $p \leq 0.01$.

CONCLUSION

The TnSeq is a powerful technique that can help to understand the mechanism behind the plant bacteria interaction. We found more than a hundred genes where the mutation is responsible for impair the epiphytic colonization of *Herbaspirillum seropedicae* in *Setaria viridis*. Between this genes we found several involved with the chemiotaxis system, indicating that this recognition and attraction between plant e bacteria is an essential step of the interaction between plant and bacteria. Several genes related with flagellum assemble were also found, but in this case, the mutation was responsible for increase the number of bacteria colonizing the tissues. The importance of the flagellum was described by many authors, but in this case the absence of this apparatus seems to be an advantage to the bacteria after 10 days of inoculation. The plants are sensitive to the Flagellum, which can be recognized by specific receptor in plant cell membrane and activate a cascade of immune responses with the goal of control the bacteria infection. Since the mutants does not have the flagellum the plant can have a delay on this recognition process and the bacteria could take an advantage and improve the colonization capacity. Another important factor during the colonization is PHB metabolism. The Carbon and energy storage also have an important role to the bacterial survival and the PHB metabolism has been shown as an essential mechanism during several situation. The function of all the proteins involved in the Synthesis e depolimerazation of PHB in *H. seropedicae* has been studied recently by several authors as Alves, 2016, Tirapelle, 2013, and Kadowaki, 2011. The PHB metabolism was also associate with the plant bacteria interaction, showing the importance of this Carbon storage pathway to the adaptation of the microbe in different environments (BALSANELLI et al., 2015 and PANKIEVICZ et al., 2016).

The importance of the PHB metabolism was described during the interaction between *H. seropedicae* and maize and wheat and now, we also indicated that the PHB metabolism are involved in the interaction with *Setaria viridis*. This shows that probably this is a pathway required for a range of different gramineas colonization, being an important pathway for the global plant bacterial interaction. Many factors can be related with the plant bacteria interaction, and many other still have to be discovery. This work gave us some insights about the pathways involved using a powerful approach and opened a variety of possibilities to continue the research focus in understand the plant bacteria interaction process.

MATERIAL AND METHODS

TnSeq Construction library

The TnSeq libraries were constructed according to Wetmore (2015). Briefly, a pKMW7 vector with a Tn5 region with different DNA bar codes was constructed. For the construction of this vector, the bar codes Amp_barcodes_FOR_Gateway_SbfI and Amp_barcodes_REV_Gateway_FseI were digested. These fragments were cloned directed into the vector pKMW4, generating the vector for the construction of libraries pKMW7. To verify the diversity of the barcodes were performed a DNA Sequencing run on MiSeq for each library. In order to avoid an overestimation of the diversity of the bar codes due to sequencing errors the minimum quality score of the sequences was 30 (Q30). Barcodes with only 1 nucleotide different from other barcodes were discarded. For the vector pKMW7, from the 13.2 million readings generated were observed 8.9 million different barcodes and 75% of the generated bar codes were identified only 1 or 2 times. The EZ:Tn5 transpososome barcode generation was done by PCR using oligonucleotides sequences complementary to the IR (inverted repeat) region of Tn5 and a random oligonucleotide with 20 base pairs complementary to the U1 and U2 sites of the BarSeq PCR priming. The EZ: TN5 bar code transpossome was prepared by addition of the PCR product, EZ: Tn5 transposase (Epicenter) and 100% glycerol. The mixture was incubated at room temperature for 30 minutes. For the construction of the mutant libraries, 1 µL of the EZ: Tn5 - Barcode transpososssoma were electroporated in electrocompetent cells of *Herbaspirillum seropedicae* SmR1. Cells from different electroporations were selected by the presence of 50 µg/mL Kanamycin. Aliquots of 1 ml of cells in OD_{600} of 1 were stored in 10% glycerol and -80 ° C and were used later for the TnSeq experiments.

TnSeq Sequencing

The sequencing was made according WETMOR (2015). Briefly, to prepare the Illumina compatible sequencing library's 1ug gDNA was fragmented in portions with the average size of 300bp by ultrasonication using a Covaris S220. The selection of the fragments with a suitable size was acquired using a AMPure XP beads (Beckman Coulter) according to the manufacturer`s instructions. The quality of the DNA was assessed in Agilent Bioanalyzer. The library preparation (end repair, A-tailing and adapter ligation) was made according with NEBNext DNA library preparation kit for Illumina (New England Biolabs). To amplify the transposons insertions sites by PCR, was used the primer Nspacer_barseq_pHIMAR

(ATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGAT CTNNNNNNCGCCCTGC AGGGATGTCCACGAG) and one primer from P7_MOD_TS_index primers containing a Illumina P7 end. For the PCR enrichment of the transposons was used a Q5 DNA polymerase with Q5 GC enhancer in the PCR program: 94°C for 2 min and 25 cycles of 94°C 30 s, 65°C for 20 s, and 72°C for 30 s, followed by a final extension at 72°C for 10 min. The PCR product was then purify using a AMPure XP beads according to the manufacturer1s instruction, eluted in water and quantify in Bioanalyzer (Agilent). The library was sequenced in either HiSeq 2000 or HiSeq 2500 system (Illumina).

TnSeq analysis

The sequencing was made according WETMORE (2015). Briefly, the reads were analyzed with a customized script (MapTnSeq.pl). The script looks for sequences U1 and U2 close to the barcode and requires a perfect match of 5 nt on each side, also requires a minimum quality score of 10 for each nt in the barcode. Only reads with at least 90% of identity and a BLAT score (matches minus mismatches) of 15 are considered. The set of barcodes that consistently map in a unique location were identify using a custom script (DesignRandomPool.pl). A barcode is considered uniquely mapped when the barcode matches with the primary location at least 10 times.

***Setaria virids* Assay**

To perform this assay, the TnSeq library was grown in LB medium in the presence of 50 μg/ml Kanamycin until it reached OD_{600} of 1. Five milliliters of the culture were collected for the sequencing of the TIME-ZERO replicates required for the analyzes. The culture was then washed 3 times with saline solution (0.9% NaCl) and resuspended in equal volume. Ten milliliters of the culture were inoculated into 5 pots containing inert soil and 15 seedlings of *Setaria viridis* A10.1 each 5 days old. The same concentration of bacteria was inoculated in 2 pots containing only soil as control. The plants were then kept in growth chambers with controlled light, humidity and photoperiod. The plants were watered each 2-3 days with Hoaglands solution (5mM of Potassium Nitrate). For the pots containing just soil was added Hoaglands with 0.05% of Malate, since the bacteria needs a minimal carbon source to growth in the inert soil. After 10 days of inoculation, the plants were recovered, the excess soil around the roots was mechanically removed and the roots were placed in culture flasks containing LB medium with Kanamycin. After 6 hours of growth the epiphytic bacteria were recovered, the LB medium filtered and the roots of Setaria viridis discarded. After 16 hours of

growth the culture reached OD_{600} of 1, the bacteria were recovered and the gDNA of the samples and of the TIME-ZERO was purified using the DNeasy Blood and Tissue kit (Quiagen).

TnSeq samples sequencing

The gDNA from the samples was quantify using a Quant-iT dsDNA BR assay kit (Invitrogen). The BarSeq PCR was performed in 50 µl of final volume, 20 µMol of each primer and 150 -200 ng of the template gDNA. The primers used was a common reverse primer (BarSeq_P1) and one of the 96 forward primers (BarSeq_P2_ITXXX) and a Q5 DNA polymerase with Q5 GC enhancer . The PCR cycle used was: 98°C for 4 min followed by 25 cycles of 30 s at 98°C, 30 s at 55°C, and 30 s at 72°C, followed by a final extension at 72°C for 5 min. Equal volumes (10 µL) of individual BarSeq PCR's were pooled and purify with the DNA Clean and Concentrator kit (Zymo Research). The BarSeq library was sequenced in either HiSeq 2000 or HiSeq 2500 system (Illumina).

BarSeq Data analysis

The analysis was made according WETMORE (2015). Briefly, the reads were converted in tables showing the number of times each barcode was seen in each sample using the custom script (MultiCodes.pl). The analysis requires a perfect match of 9 nucleotides upstream of the barcode, otherwise the read is discarded. Knowing the table of barcodes, where they map in the genome and their counts in each sample was estimate the strains fitness and gene fitness values with the custom R script FEBA.R. The strain fitness is the normalized log2 ratio of counts between the treatment sample and the reference "time-zero" sample. The gene fitness is the weighted average of the strain fitness. All the experiments were made with 2-5 replicates with independent gDNA extraction and independent PCR. Gene fitness ≥ 1 or ≤ -1 was considered different. For the statistical analysis was used a moderated *t* statistic, genes with $|t|$ of 4 have highly significant phenotypes that are reproducible in biological replicate experiments (see WETMORE et al., 2015).

Plant-bacteria assay

The Setaria viridis A10.1 was used for this experiments. The *Setaria viridis* seeds were sterilized for 5 minutes in a solution of 5% sodium hypochlorite and 0.2% of tween 20%. The seed were then washed at least 3 times with sterile MilliQ water for 2 minute. Seeds were germinated in 1% Plant Medium plates for 48 in the dark and 24 hours in light - 30°C. The

seedlings were inoculated with bacterial suspensions containing 10^5 bacteria/mL for 30 minutes. The seedlings were then transferred to experimental tubes containing 20mL of Plant Medium (EGENER et al., 1999) solution, with Polypropylene spheres and kept at 25°C with a 12 hours of photoperiod cycle. To determine the total number of root attached bacteria, the plants were recovered, washed 3 times with a saline (NaCl 0,9%) solution, heavily shaken to release the epiphytic bacteria and the supernatant was serially diluted and plated on NFbHPN malate medium.

AUTHOR CONTRIBUTION

Adan Deutschbauer (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA) – Assisted the TnSeq library construction, sequencing and data analysis

Adan Arkin (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA and Department of Bioengineering, University of California, Berkeley, California, USA) - Assisted the TnSeq library construction, sequencing and data analysis

Fernanda P do Amaral and Beverly Jose Agtuca (Division of Plant Science and Biochemistry, University of Missouri, USA) - Assisted the Sorghum Screening

Tomas P Pereira (Department of Science and Food Technology, Federal University of Santa Catarina, Brazil) - Assisted the Sorghum Screening

Fabio de Oliveira Pedrosa, Emanuel Maltempi de Souza and Rose Adele Monteiro (Department of Biochemistry and Molecular Biology, Federal University of Paraná, Brazil) – Orientation and Supervision

Gary Stacey (Division of Plant Science and Biochemistry, University of Missouri, USA) - Orientation and Supervision

REFERENCES

ALVES, L.P.S., TEIXEIRA, C. S., TIRAPELLE, E.F., DONATTI, L., TADRA-SFEIR, M. Z., STEFFENS, M. B. R., SOUZA, E.M., PEDROSA, F. O., CHUBATSU, L. S., MÜLLER-SANTOS, M. Backup Expression of the PhaP2 Phasin Compensates for phaP1 Deletion in *Herbaspirillum* seropedicae, Maintaining Fitness and PHB Accumulation. **Front. Microbiol**. 7:739. 2016.

BABEL, W., ACKERMANN, J. U., AND BREUER, U. Physiology, regulation, and limitsofthesynthesisofpoly(3HB).Adv.**Biochem.Eng.Biotechnol**.71,125–157. 2001

BALDANI, J.I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum* seropedicae gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. **International Journal of Systematic Bacteriology**. V. 36. P. 86-93. 1986.

BALDANI, V.L.D.; BALDANI, J.I.; OLIVARES, F.; DÖBEREINER, J. Identification and ecology of *Herbaspirillum* seropedicae and the closely related Pseudomonas rubrisubalbicans. **Symbiosis**, v. 13, p. 65-73, 1992.

BALDANI, V.L.D.; BALDANI, J.I.; DOBEREINER, J. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum* seropedicae and Burkholderia spp. **Biol. Fertil. Soils** 30:485-491, 2000.

BALSANELLI E, TULESKI TR, DE BAURA VA, YATES MG, CHUBATSU LS, PEDROSA, F. O., SOUZA, E. M., MONTEIRO, R. A. Maize Root Lectins Mediate the Interaction with *Herbaspirillum* seropedicae via N-Acetyl Glucosamine Residues of Lipopolysaccharides. **PLoS ONE.** 2013.

BALSANELLI, E., TADRA-SFEIR, M. Z., FAORO, H., PANKIEVICZ, V., BAURA, V. A., PEDROSA, F. O., SOUZA, E. M. DIXON, R. & MONTEIRO, R. A. Molecular adaptations of *Herbaspirillum* seropedicae during colonization of the maize rhizosphere. **Environmental microbiology**. 2015.

BERG, H. C. The rotary motor of bacterial flagella. **Annu. Rev. Biochem**. 72, 19-54 2003.

BERG G, EBERL L, HARTMANN A (2005) The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. **Environ Microbiol** 7:1673–1685

BERG, H.C., BROWN, D. A. Chemotaxis in Escherichia coli analysed by three-dimensional tracking. **Nature**, 239:500–4. 1972

BODDEY, R.M., DE OLIVEIRA, O.C., URQUIAGA, S., REIS, V.M., OLIVARES, F.L., ET AL. Biological nitrogen fixation associated with sugar cane and rice: contributions and prospects for improvement. **Plant Soil** 174: 195–209. 1995

BROEK, A. V., LAMBRECHT, M., VANDERLEYDEN, J. Bacterial chemotactic motility is important for the initiation of wheat root colonization by Azospirillum brasilense. **Microbiology.** V144, 1998.

CORDEIRO, F. A., TADRA-SFEIR, M. Z., HUERGO, L. F., PEDROSA, F. O., MONTEIRO, R. A., AND SOUZA, E. M. Proteomic analysis of *Herbaspirillum* seropedicae cultivated in the presence of sugar cane extract. **J. Proteome Res**. 12, 1142–1150. 2013.

DONG, T.G., HO, B.T., YODER-HIMES, D.R., AND MEKALANOS, J.J. Identification of T6SS-dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. **Proc Natl Acad Sci USA** 110: 2623–2628. 2013.

EGENER, T., HUREK, T., AND REINHOLD-HUREK, B. Endophytic expression of nif genes of Azoarcus sp. strain BH72 in rice roots. **Mol Plant Microbe Interact** 12: 813–819. 1999.

ELBELTAGY, A., NISHIOKA, K., SATO, T., SUZUKI, H., YE, B., HAMADA, T., et al. Endophytic colonization and in planta nitrogen fixation by a *Herbaspirillum* sp. isolated from wild rice species. **Appl Environ Microbiol** 67: 5285– 5293. 2001

GYANESHWAR, P., JAMES, E.K., REDDY, P.M., AND LADHA, J.K. *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. **New Phytol** 154: 131–145. 2002.

HERVAS, A. B., CANOSA, I., AND SANTERO, E. (2008). Transcriptome analysis of Pseudomonas putida in response to nitrogen availability. **J. Bacteriol**. 190, 416–420.doi:10.1128/JB.01230-07

INIGUEZ AL, DONG Y, TRIPLETT EW (2004) Nitrogen fixation in wheat provided by Klebsiella pneumoniae 342. **Mol Plant Microbe Interact** 17:1078–1085
JAMES, E.K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crop Res** 65: 197–209. 2000.

JAMES, E.K., GYANESWAR, P., MATHAN, N., BARRAQUIO, W.L., REDDY, P.M., IANNETTA, P.P.M. Infection and colonization of the rice seedlings by the plant growthpromoting bacterium *Herbaspirillum* seropedicae Z67. **Mol Plant Microbe Interact** 15: 894–906. 2002.

JURASEK, L., AND MARCHESSAULT, R. H. The role of phasins in the morphogenesis of poly(3-hydroxybutyrate) granules. **Biomacromolecules** 3, 256. 2002.

KADOWAKI, M.A., MULLER-SANTOS, M., REGO, F.G., SOUZA, E.M., YATES, M.G., et al. (2011) Identification and characterization of PhbF: a DNA binding protein with regulatory role in the PHB metabolism of *Herbaspirillum* seropedicae SmR1. **BMC Microbiol** 11: 230. doi:10.1186/1471-2180-11-230.

LIU, X., MATSUMURA, P. The FlhD/FlhC Complex, a Transcriptional Activator of the Escherichia coli Flagellar Class II Operons. **Journal of bacteriology**, Dec. 1994, p. 7345-7351 Vol. 176, No. 23

MONTEIRO, R.A., BALSANELLI, E., WASSEM, R. et al. *Herbaspirillum*-plant interactions: microscopical, histological and molecular aspects**. Plant Soil** (2012) 356: 175.

NEUMANN, L., SPINOZZI, F., SINIBALDI, R., RUSTICHELLI, F., POTTER, M., AND STEINBUCHEL, A. (2008). BINDING OF THE MAJOR PHASIN, PHAP1, FROM RALSTONIA EUTROPHA H16 TO Poly(3-hydroxybutyrate) granules. **J. Bacteriol**. 190, 2911– 2919.doi:10.1128/JB.01486-07

OPIJNEN, T. V. AND CAMILLI, A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. **Nat Rev Microbiol** . 2013 July ; 11(7).

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., SOUZA E. M. RNA-seq transcriptional profiling of *Herbaspirillum* seropedicae colonizing wheat (Triticum aestivum) roots. **Plant Mol Biol**, 2016.

PEDROSA, F.O.; MONTEIRO, R. A.; WASSEM, R.; CRUZ, L. M. Genome of *Herbaspirillum* seropedicae strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**. doi:10.1371/journal.pgen.1002064. 2011.

RONCATO-MACCARI, L.D.B., RAMOS, H.J.O., PEDROSA, F.O., ALQUINI, Y., CHUBATSU, L.S., YATES, M.G. Endophytic *Herbaspirillum* seropedicae expresses nif genes in gramineous plants. **FEMS Microbiol Ecol** 45: 39–47. 2003.

SESSITSCH, A.; HOWIESON, J. G.; PERRET, X.; ANTOUN, H.; MARTÍNEZ-ROMERO, E. Advances in rhizobium research. **Crit Ver Plant Sci** 21:323–378. 2002

STURZ, A. V., CHRISTIE, B. R., NOWAK, J. Bacterial endophytes: potential role in developing sustainable systems of crop production. **CRC Crit Rev Plant Sci** 19:1–30 2000.

TADRA-SFEIR, M. Z., FAORO, H., CAMILIOS-NETO, D., BRUSAMARELLO-SANTOS, L., BALSANELLI, E., WEISS, V., BAURA, V. A., WASSEM, R., CRUZ, L. M., PEDROSA, F. O., SOUZA, E. M., MONTEIRO, R. A. Genome wide transcriptional profiling of *Herbaspirillum* seropedicae SmR1 grown in the presence of naringenin. **Frontiers in Microbiology**. 6:491. 2015

TIRAPELLE, E.F., MÜLLER-SANTOS, M., TADRA-SFEIR, M.Z., KADOWAKI, M.A.S., STEFFENS, M.B.R., et al. (2013) Identification of proteins associated with polyhydroxybutyrate granules from *Herbaspirillum* seropedicae SmR1 – old partners, new players. **PLoS ONE** 8: e75066.

WETMORE KM, PRICE MN, WATERS RJ, LAMSON JS, HE J, HOOVER CA, BLOW MJ, BRISTOW J, BUTLAND G, ARKIN AP, DEUTSCHBAUER A. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. **mBio** 6(3). 2015.

CAPÍTULO VI

Manuscrito ainda não submetido

Identification of essential *Herbaspirillum seropedicae* genes in different growing conditions using RB-TnSeq

Identification of essential *Herbaspirillum seropedicae* genes in different growing conditions using RB-TnSeq

ABSTRACT

The *Herbaspirillum seropedicae* is a nitrogen fixing bacteria found in association with many important crops. Several works has been published aiming the better understand the metabolic pathways of this microbe. The TnSeq is a new approach, recently used for the study of plant associative bacteria and has been demonstrated as good strategy to perform gene function. In this work we challenge the *Herbaspirillum seropedicae* SmR1 in different growth conditions, as different carbon and nitrogen source and osmotic stress adaptation. We could identify several genes involved with the environmental conditions. A large number of transporters and transcriptional regulators involved specifically with the analyzed conditions were identified. A large number of hypothetical proteins was also found, suggesting that this technique could be a good tool to annotate several genes of *Herbaspirillum seropedicae*.

INTRODUCTION

Endophytic diazotrophic bacteria are organisms able to associate and promote a beneficial interaction with several plants, fixing the atmospheric nitrogen and making it available to the plants. Among the known nitrogen-fixing organisms are some species of the Betaproteobacteria genus *Herbaspirillum*, such as *Herbaspirillum seropedicae* (BALDANI et al., 1986; OLIVARES et al., 1997).

The mechanisms of association between *H. seropedicae* and plants is not fully understood. It is already known that they start with the adhesion of the bacterium to the root surface, colonization of the secondary root points and discontinuities of the epidermis followed by penetration and scattering of the bacteria in the intercellular spaces of roots, aerenchyma, vascular tissues and aerial portions (JAMES et al. , 1997; RONCATO-MACCARI et al., 2003).

H. seropedicae can be associated with several plants of economic interest, such as maize (*Zea mays*), rice (*Oryza sativa*), sorghum (*Sorghum bicolor*) and sugarcane (*Saccharum officiarum*) (BALDANI et al., 1986 BALDANI et al., 1992). This interaction is beneficial for both organisms, and may increase the accumulated nitrogen and promote plant growth

(BALDANI et al., 2000 and PANKIEVICZ et al., 2015), indicating the great potential of these bacteria as a nitrogen biofertilizer.

To the total understanding of the biology of this organism it is important as the elucidation of all pathways and requirements for growth of this organism. Several studies about the Carbon requirements and metabolism have been published with the goal of totally understand the metabolism of *H. seropedicae* SmR1.

The DNA Barcode Transposon Sequencing (RB-TnSeq) was recently used for bacteria and it is a powerful methodology to the global study of bacterial requirement and gene annotation. The BarSeq (RB-TnSeq) is a transposon sequencing experiment, using DNA BarCodes to identify the mutation and determine the fitness profile of the strains. The recent description of this work by Wetmore (2015) shows a high efficiency and reproducibility of the results in an experiment using more than a hundred different Carbon sources in 5 microbe species: *Escherichia coli, Phaeobacter inhibens, Pseudomonas stutzeri, Shewanella amazonensis, and Shewanella oneidensis.* This work could identify more than 5 thousand different important genes dependent of the Carbon source in this 5 bacteria using the mutant fitness profile. Here we focus in elucidate the essential genes and pathways for the *Herbaspirillum seropedicae* growth in different sources of Carbon, Nitrogen and Osmotic Stress using the RB-TnSeq technique (WETMORE et al., 2015).

RESULTS AND DISCUSSION

A TnSeq assay was performed with *Herbaspirillum seropedicae* SmR1 in NFb medium with different sources of Carbon, Nitrogen and Osmotic Stress (TABLE 6-1). For the assay, the TnSeq library was grown in LB medium plus Kanamycin until it reached OD_{600} of 1. The bacteria were then washed three times with saline solution and reinoculated in the media of interest at the initial OD_{600} of 0.02. Cultures were maintained at 30° C and 120 rpm until they reached OD_{600} of 1 and then were recovered and their gDNA extracted using the DNeasy Blood and Tissue kit.

TABLE 6-1 – Carbon, Nitrogen and Osmotic Stress Sources used in the RB-TnSeq assay

Carbon Source	**Concentration**	
D-galactose	5 g/L	NH_4Cl – 20 mM
D-glucose	5 g/L	NH_4Cl – 20 mM
DL-sodium malate	5 g/L	NH_4Cl – 20 mM
D-xylose	5 g/L	NH_4Cl – 20 mM
L-rhamnose	5 g/L	NH_4Cl – 20 mM
L-fucose	5 g/L	NH_4Cl – 20 mM
D-mannose	5 g/L	NH_4Cl – 20 mM
D-mannitol	5 g/L	NH_4Cl – 20 mM
DL-Lactate	5 g/L	NH_4Cl – 20 mM
Nitrogen Source		
Ammonium Chloride	1 mM	NFbHP Malate – HIGH AERATION
Ammonium Chloride	1mM	NFbHP Malate – LOW AERATION
L-Glutamate	1 mM	NFbHP Malate - HIGH AERATION
L-Glutamate	1 mM	NFbHP Malate - LOW AERATION
$NaNO_3$	10 mM	NFbHP Malate - HIGH AERATION
NaNO3	10 mM	NFbHP Malate - LOW AERATION
Osmotic Stress		
NaCl	1%	NFbHP Malate + 20mM de NH_4Cl

SOURCE: The Author (2017).

After the gDNA extraction the sequencing was performed in Illumina HiSeq platform. The fitness for each gene was generate using the FEBA.R package (R program). The genes with the *gene fitness* $\leq$ -1 and $\geq$ 1 and with statically t test $\geq |4|$ were considerated as essential genes. The strains with positive *fitness* are considered with Beneficial mutation, when the mutation increase the ability of the bacteria growth in the determined condition. The genes with negative *fitness* are considered as Sick Mutations, when the mutation decreases the ability of the bacteria growth in the determined condition.

Carbon Source Essential Genes

For the Carbon requirement analysis, the bacteria were grown in different Carbon Sources. Before the inoculation, the bacteria were washed 3 times to avoid another carbon sources from the pre-inoculum media. The bacteria were inoculated in NFb media with 20 mM of ammonium chloride as nitrogen source. The number of genes found in each condition is shown in the table 6-2.

TABLE 6-2 – Number of genes with fitness profile significantly different from the TIME-ZERO in each Carbon condition. (The data is shown as Number of exclusive genes in the condition/number of total genes in the condition).

	Fucose	Galactose	Glucose	Lactate	Malate	Mannitol	Mannose	Rhamnose	Xylose
Sick	9/84	6/98	3/114	28/124	16/115	12/112	32/153	12/121	14/116
Beneficial	25/27	0	3/4	3/7	0/1	0/2	17/22	1/4	1/3
Total	**34/111**	**6/98**	**6/118**	**31/131**	**16/116**	**12/114**	**49/175**	**13/125**	**15/119**

SOURCE: The Author (2017).

A comparison was made between all treatments with the goal of finding genes present in all conditions (genes essential for general carbon metabolism) and genes that are unique to each condition, being determined by the type of carbon source available.

The genes were categorized according to the Clusters of Orthologous Groups of proteins (COG). Between the genes shared by all conditions, the most part are related to the amino acid metabolism and transport (Class E – COG). The second class more abundant was P (Inorganic Ion Transport and metabolism), with 4 genes; proteins in the COG classes C, D, F, H, L, M, and S were also found in the analysis (FIGURE 6-1).

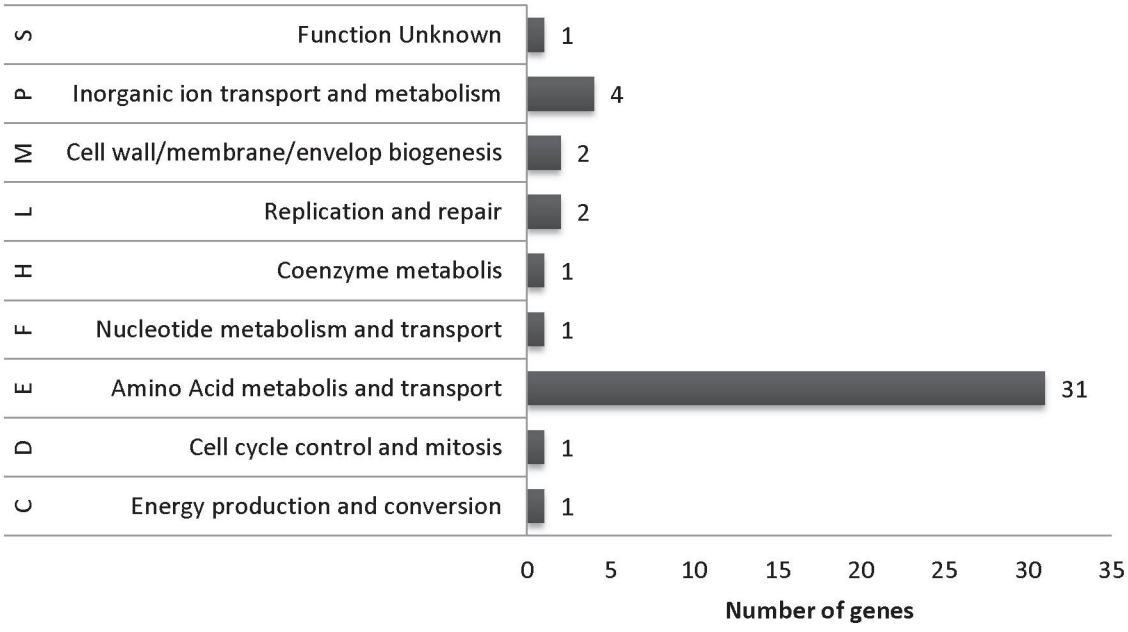


FIGURE 6-1 – COG (Clusters of Orthologous Groups of proteins) classification of the common essential proteins for all the Carbon Sources.

Malate as Carbon Source

When the DL-sodium malate was the Carbon Source, we found 16 genes of *H. seropedicae* exclusively present NFbHP with DL-sodium malate. The protein (HSERO_RS07390) tartrate dehydrogenase (TDH) was identified and the gene fitness of this mutant was -3, suggesting that this protein has an important role in the Malate metabolism. The TDH is a NAD-dependent enzyme with several catalytic activities and found in a diversity of microorganisms. One of their functions is convert tartrate into D-glycerate, promoting the utilization of this C atoms in the primary metabolic pathways (TIPTON et al., 2000; TIPTON and PEISACH, 1990). This enzyme also can catalyze the oxidation of (+)-tartrate into oxaloglycolate and it is capable of catalyze the oxidative decarboxylation of D-malate, producing pyruvate and CO_2 (MALIK et al., 2010) (FIGURE 6-2). The fitness of -3 suggest that this enzyme is crucial for the growth in this media.

We also found a LysR family transcriptional regulator (HSERO_RS07400) with a gene fitness of -2.6, suggesting that this regulator is involved with the malate assimilation. This family comprised several autoregulatory transcriptional factors which could be involved with extremely diverse functions (SHELL et al., 1993). This regulator are genetically close to an aldose-1-epimerase, suggesting that it may have a function in the Glycolisis.

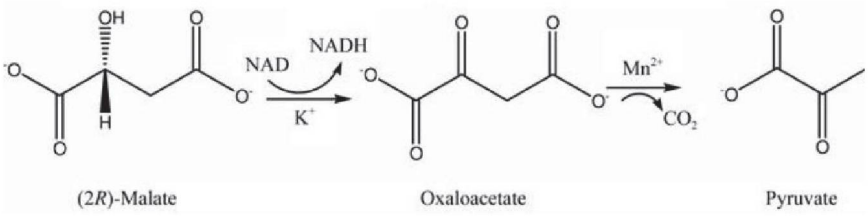


FIGURE 6-2 – Conversion of Malate into Pyruvate pathway catalyzed by tartrate dehydrogenase (TDH)

Galactose as Carbon source

For the galactose condition, we found 6 exclusive genes, between then 2 are involved directly with the galactose assimilation pathway: *dgoAa* (2-dehydro-3-deoxy-6-phosphogalactonate aldolase, HSERO_RS05155, *gene fitness*=-3) and *dgoK* (2-dehydro-3-deoxygalactonokinase, HSERO_RS05160, *gene fitness*=-5.7). Both this genes are strictly necessary for the growth in galactose, being the *dgoK* mutant strain incapable of growth in a media with Galactose as unique Carbon Source (FIGURE 6-3).

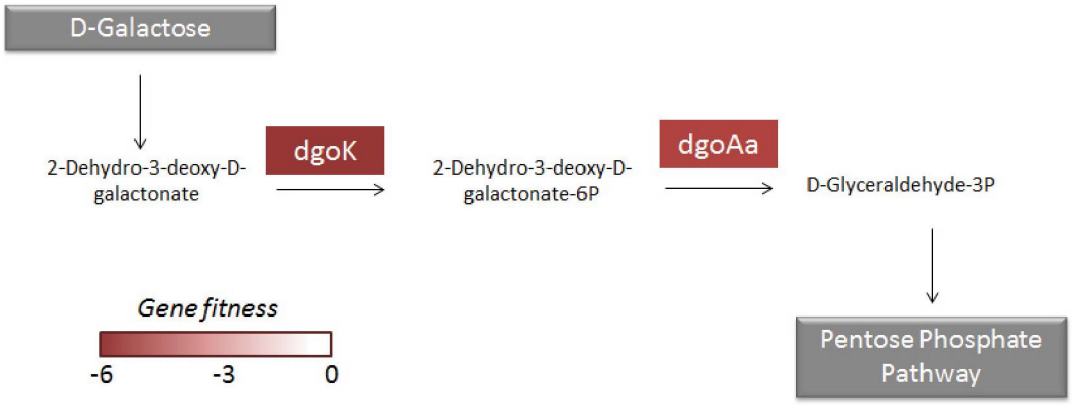


FIGURE 6-3 – Galactose assimilation pathway. The enzymes DgoK and DgoAa are essential for the growth in Galactose media, being the *dgoK* mutant incapable to growth in this media.

We also found 2 transcriptional regulator HSERO_RS09235 (MarR family transcriptional regulator) and HSERO_RS21480, a phosphoheptose isomerase (HSERO_RS00410) and a hypothetical protein (HSERO_RS13880); all of them has gene fitness around 1 suggesting that they are important but not essential for the growth in galactose.

Mannitol as Carbon source

During the mannitol growth we could find 12 exclusive genes, between them we found 4 ABC sugar transporter ATP-binding protein: SmoK (HSERO_RS02210), SmoF(HSERO_RS02195), SmoG (Hsero_0442) and SmoE (Hsero_0440). All of them has a gene fitness between -3 and -3.7, suggesting that the entry of mannitol into the cell it is a crucial step for the mannitol metabolism.

Mannose as Carbon source

We found 5 sugar transporters exclusively important for the growth in mannose: D-ribose transporter ATP binding (HSERO_RS03640, *gene fitness* -4.3), ribose ABC transporter permease (HSERO_RS03645, *gene fitness -2.9*), ABC transporter (HSERO_RS10850, *gene fitness -2.1*), sugar ABC transporter periplasmic protein (HSERO_RS03635, gene fitness -4.7) and a fructose transporter (HSERO_RS03630, gene fitness -4.8). The profile fitness of all

transporters suggest that they are crucial for the mannose assimilation. We also found a N-acyl-D-glucosamine 2-epimerase (HSERO_RS03650) with the gene fitness of -5.4, this enzyme is responsible for catalyze the conversion of N-acyl-D-glucosamine 6-phosphate into N-acyl-D-mannosamine 6-phosphate and based on the fitness profile, this is an important enzyme for the mannose assimilation.

Rhamnose as Carbon source

For the Rhamnose analysis, we identified 12 exclusive genes including the following sugar transporters: RhaT D-ribose transporter (HSERO_RS22220, gene fitness -3.5), RhaO (HSERO_RS22210 branched –chain amiacid ABC transporter permease, gene fitness -4.8), RhaP (HSERO_RS22215 - branched –chain amiacid ABC transporter permease, gene fitness -4.2), RhaS (HSERO_RS22225, ABC transporter substratate-binding protein, gene fitness -4.6). We identified a sugar isomerase, RhaI (HSERO_RS22240, gene fitness -4.6) responsible for the interconversion of L-Rhamnose to L-Rhamnulose. Another important protein found was an L-Rhamnose mutarotase (HSERO_RS22205, gene fitness -2.5), this family of proteins is responsible for interconvert the alpha and beta stereoisomers of monosaccharides when the less-favored anomer is necessary for next step in a catabolic pathway (RICHARDSON et al., 2008). A short-chain dehydrogenase (HSERO_RS22235) was found with very critical gene fitness of -6. This enzymes are from a large family of NAD(P)(H)-dependent oxidoreductases and have a critical role in the carbohydrate, lipid, amino acid, cofactor, hormone and xenobiotic metabolism as well as in a redox sensor mechanisms (KAVANAGH et al., 2008). Based on the very critical *gene fitness* profile, this enzyme may have a crucial function during the assimilation of Rhamnose. Another important proteins indicated by the gene fitness are a transcriptional regulator from DeoR family (HSERO_RS22230) with the gene fitness of -6 and a L-fuculose kinase (HSERO_RS22200) with the gene fitness of -5.7.

General analysis

Between all the genes that are specific for each carbon source, the most part comprise the class G of the COG: Carbohydrate transport and Metabolism, as expected (FIGURE 6-4). If we compare with COG classification of the genes present in all conditions at the same time, we could not find any gene related to the this category, showing that this proteins are exclusively and dependent of the carbon source utilized. The second most abundant COG

category is the K (Transcription) which includes several transcriptional regulators. This analysis suggest that this process are strictly regulated, since the carbon metabolism pathways involve a high amount of energy production and conversion is expected that the bacteria has a fine regulation in this growth conditions. The third most abundant COG category is S, Unknown Function, suggesting that still there are many genes involved with this pathways should be elucidated. Other categories, such as Energy Production and Coenzyme transport and metabolism also has a good representativeness in the analysis. The table with all exclusive genes found for each C source is shown in Supplementary table 6-1.

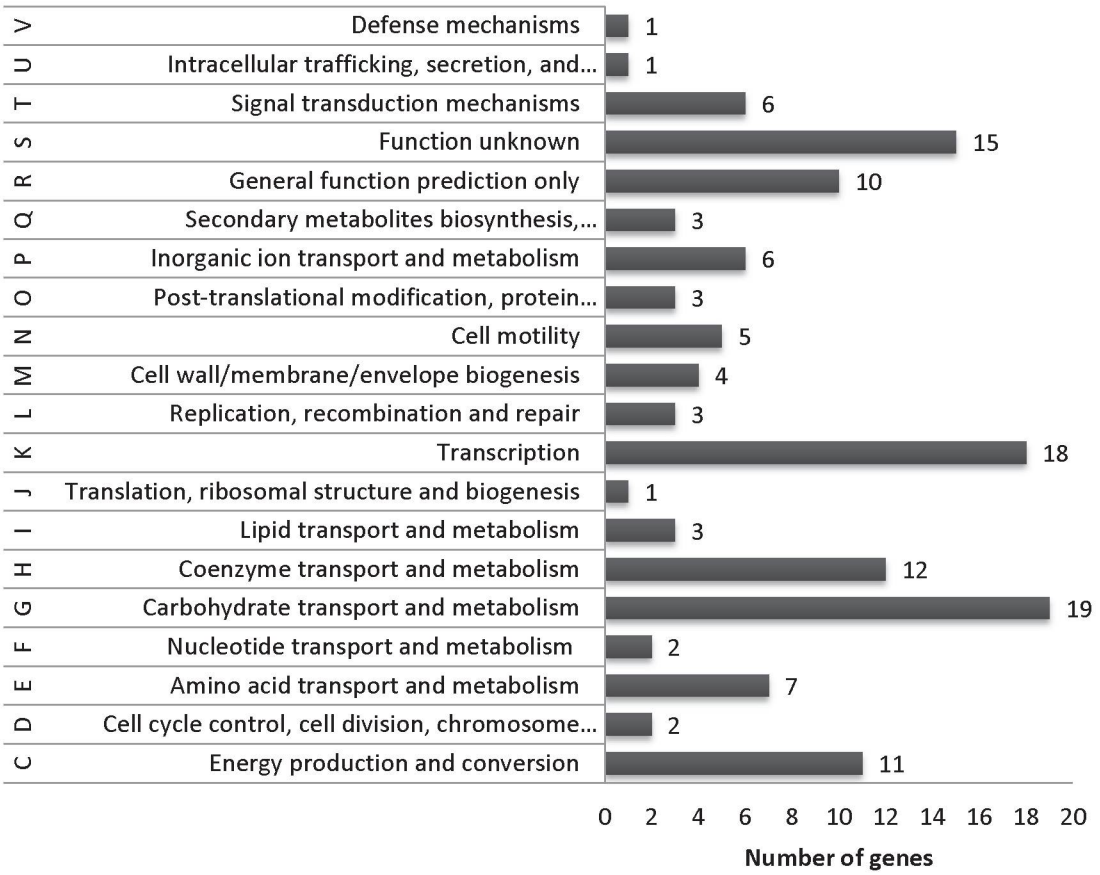


FIGURE 6-4 - COG (Clusters of Orthologous Groups of proteins) classification of the specific essential proteins for each Carbon Sources.

Osmotic Stress

A TnSeq assay was also performed in the presence of 1% NaCl (osmotic stress). The TnSeq library was grown on NFbHP Malate medium, with 20 mM NH_4Cl as nitrogen source.

We found 129 SICK genes and 13 genes where the mutation is beneficial for growth. Among the SICK genes, 90 were also found in the control treatment (NFbHP Malate and 20 mM NH_4Cl) and 39 genes are unique in the growth in NaCl.

Between the genes exclusively found in the NaCl condition we found a Cation transporter (HSERO_RS03595) and two C4-dicarboxylate ABC transporters (HSERO_RS13105 and HSERO_RS04235, with gene fitness of -2.9 and -2.1 respectively). These transporters are part of a family of transmembrane ion transporters, responsible for exchange dicarboxylates, as malate and succinate. The Rb-TnSeq analysis suggests that these proteins could be involved with the bacterial response against the Osmotic Stress. The table with all genes identified during the Osmotic stress is shown in Supplementary table 2.

Nitrogen Metabolism

The genes involved in Nitrogen Metabolism were also analyzed in different conditions. We perform an assay using 3 different Nitrogen sources and different levels of aeration. The levels of aeration were adjusted based on the air column in the flask, having the Low aeration (LA) 66% of air in the flask and the High Aeration (HA) 95% of air in the flask. The number of genes where the mutation affects the fitness of the bacteria during the Nitrogen Source variation and aeration is shown in Table 6-4.

TABLE 6-3 – Number of genes with fitness profile significantly different from the TIME-ZERO in each Nitrogen condition

	Nitrate Low Aeration	**Nitrate High Aeration**	**NH_4Cl Low Aeration**	**NH_4Cl High Aeration**	**Glutamate Low Aeration**	**Glutamate High Aeration**
Sick	132	133	77	99	78	91
Beneficial	21	1	3	1	2	5
Total	**153**	**134**	**80**	**100**	**80**	**96**

SOURCE: The Author (2017).

Venn diagrams (FIGURE 6-5) show specific genes found essentials related to different nitrogen sources and under high and low aeration growth condition.

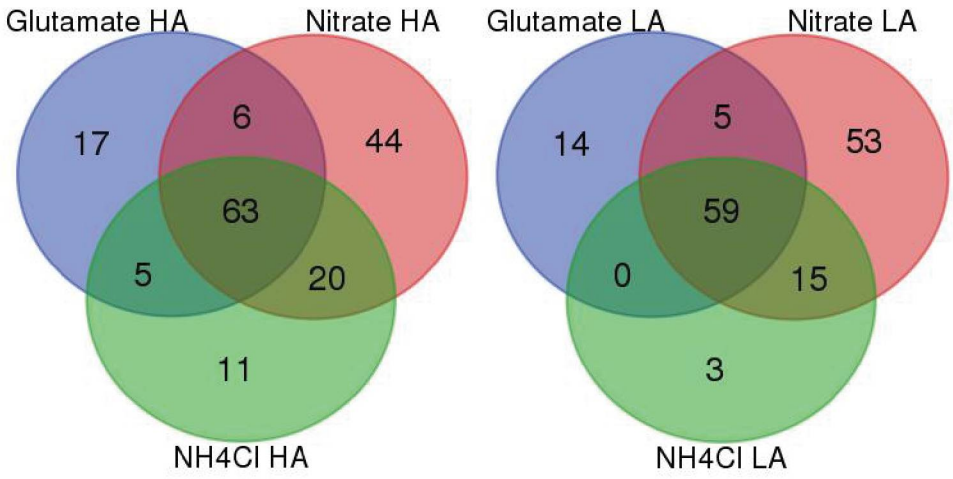


FIGURE 6-5 - Number of Sick genes where the mutation was important for growth under the indicated conditions (HA = High aeration and LA = Low aeration).

When the Nitrogen source used was Glutamate in both high and low aeration we could identify 2 essential glutamate transporters, as expected: gltL - arginine ABC transporter ATP-binding protein (HSERO_RS19240) and gltJ - glutamate ABC transporter permease (HSERO_RS19250). The ABC Transporter (HSERO_RS19255) was also found in high aeration.

When the Nitrogen source is the Nitrate in both high and low aeration we identify the genes *nirD* – nitrite reductase (HSERO_RS14540, gene fitness -4) and *narK* - MFS transporter (HSERO_RS14530, gene fitness -2.2). The *nasA* - nitrate reductase (HSERO_RS14550, gene fitness -2) was found in high aeration while the *nirB* – nitrite reductase (HSERO_RS14535, gene fitness -4.1) was identified in Low aeration. These proteins were already described as important components of the Nitrate metabolism in *Herbaspirillum seropedicae* SmR1, the NarK is a nitrate transporter, responsible for translocate the nitrate from the extracellular environment to the intracellular media, where the nitrate can be reduced to nitrite by the NasA, an assimilatory nitrate reductase. The nitrite generated will be reduced by a nitrite reductase, NirBD, into ammonium which will utilize in the molecules structures (PEDROSA et al., 2011; BONATO et al., 2016). The expression of the nitrite reductase in low oxygen concentration was demonstrated in *E. coli* with the function of nitrite cell detoxification after the nitrate assimilation (MORENO-VIVIÁN et al., 1999), this data support our data, being the *nirB* just found in low aeration, with a critical

fitness profile indicating this enzyme essential for the *H. seropedicae* growth in nitrate under low aeration.

The table with the genes and their gene fitness in High Aeration when we compare the Nitrogen Sources are shown in the Supplementary Table 3, the data related to the Low Aeration when we compare the Nitrogen Sources is shown in the Supplementary Table 4.

The tables comparing the same Nitrogen Source in Low and High Aeration is shown in the Supplementary Table 5 for Glutamate, Table 6 for Nitrate and table 7 for Ammonium chloride.

CONCLUSIONS

The TnSeq analysis is a powerful approach to analyze the essential genes for different conditions. In this work we did a screening of essential genes for *Herbaspirillum seropedicae* SmR1 in different conditions of Carbon, Nitrogen and osmotic stress sources. We could identify a large number of candidates genes potentially involved specifically with the conditions analyzed. Several specific transporters were identified, as expected, as well as diverse group of transcriptional regulators. A large number of unknown proteins were also identified, suggesting that this technique could be extremely useful for the gene annotation.

MATERIAL AND METHODS

TnSeq Construction library

The TnSeq libraries were constructed according to Wetmore (2015). Briefly, a pKMW7 vector with a Tn5 region with different DNA bar codes was constructed. For the construction of this vector, the bar codes Amp_barcodes_FOR_Gateway_SbfI and Amp_barcodes_REV_Gateway_FseI were digested. These fragments were cloned directed into the vector pKMW4, generating the vector for the construction of libraries pKMW7. To verify the diversity of the barcodes were performed a DNA Sequencing run on MiSeq for each library. In order to avoid an overestimation of the diversity of the bar codes due to sequencing errors the minimum quality score of the sequences was 30 (Q30). Barcodes with only 1 nucleotide different from other barcodes were discarded. For the vector pKMW7, from the 13.2 million readings generated were observed 8.9 million different barcodes and 75% of the generated bar codes were identified only 1 or 2 times. The EZ:Tn5 transpososome barcode generation was done by PCR using oligonucleotides sequences complementary to the IR

(inverted repeat) region of Tn5 and a random oligonucleotide with 20 base pairs complementary to the U1 and U2 sites of the BarSeq PCR priming. The EZ: TN5 bar code transpossome was prepared by addition of the PCR product, EZ: Tn5 transposase (Epicenter) and 100% glycerol. The mixture was incubated at room temperature for 30 minutes. For the construction of the mutant libraries, 1 µL of the EZ: Tn5 - Barcode transpososssoma were electroporated in electrocompetent cells of *Herbaspirillum seropedicae* SmR1. Cells from different electroporations were selected by the presence of 50 µg/mL Kanamycin. Aliquots of 1 ml of cells in OD_{600} of 1 were stored in 10% glycerol and -80 ° C and were used later for the TnSeq experiments.

TnSeq Sequencing

The sequencing was made according WETMORE (2015). Briefly, to prepare the Illumina compatible sequencing library's 1ug gDNA was fragmented in portions with the average size of 300bp by ultrasonication using a Covaris S220. The selection of the fragments with a suitable size was acquired using a AMPure XP beads (Beckman Coulter) according to the manufacturer`s instructions. The quality of the DNA was assessed in Agilent Bioanalyzer. The library preparation (end repair, A-tailing and adapter ligation) was made according with NEBNext DNA library preparation kit for Illumina (New England Biolabs). To amplify the transposons insertions sites by PCR, was used the primer Nspacer_barseq_pHIMAR (ATGATACGGCGACCACCGAGATCTACACTCTTTC CCTAC ACGACGCTCTTCCGA TCTNNNNNNCGCCCTGCAGGG ATGTCCACG AG) and one primer from P7_MOD_TS_index primers containing a Illumina P7 end. For the PCR enrichment of the transposons was used a Q5 DNA polymerase with Q5 GC enhancer in the PCR program: 94°C for 2 min and 25 cycles of 94°C 30 s, 65°C for 20 s, and 72°C for 30 s, followed by a final extensionat72°C for 10 min. The PCR product was then purify using a AMPure XP beads according to the manufacturer1s instruction, eluted in water and quantify in Bioanalyzer (Agilent). The library was sequenced in either HiSeq 2000 or HiSeq 2500 system (Illumina).

TnSeq analysis

The reads were analyzed with a customized script (MapTnSeq.pl). The script looks for sequences U1 and U2 close to the barcode and requires a perfect match of 5 nt on each side, also requires a minimum quality score of 10 for each nt in the barcode. Only reads with at least 90% of identity and a BLAT score (matches minus mismatches) of 15 are considered. The set of barcodes that consistently map in a unique location were identify using a custom

script (DesignRandomPool.pl). A barcode is considered uniquely mapped when the barcode matches with the primary location at least 10 times.

Bacterial Growth Conditions

To perform this assay, the TnSeq library was grown in LB medium in the presence of 50 μg/ml Kanamycin until it reached OD_{600} of 1. Five milliliters of the culture were collected for the sequencing of the TIME-ZERO replicates required for the analyzes. The bacteria was washed 3 times with saline solution (NaCl 0.9%) and then were reinoculated in the appropriated media in an initial OD_{600} of 0.02. Cultures were grown until the bacteria reached OD_{600} of 1 or for at least 4 generations, at 30°C and 120 rpm. For the samples with low and high aeration, the air column in the flask was modified. For the high aeration, a 20 mL flask with 3 mL of bacterial culture was used. For the low aeration, the same flask of 20 mL was used to growth 10 mL of bacterial culture. The total genomic DNA was extracted using a DNeasy Blood and Tissue Kit - Quiagen. For all the experiments in different Carbon sources the media NFbHP with 20mM of NH4Cl was used (KLASSEN et al., 1997). The bacteria was grown at 30°C and 120 rpm.

TnSeq samples sequencing

The gDNA from the samples was quantify using a Quant-iT dsDNA BR assay kit (Invitrogen). The BarSeq PCR was performed in 50ul of final volume, 20uMol of each primer and 150 -200 ng of the template gDNA. The primers used was a common reverse primer (BarSeq_P1) and one of the 96 forward primers (BarSeq_P2_ITXXX) and a Q5 DNA polymerase with Q5 GC enhancer . The PCR cycle used was: 98°C for 4 min followed by 25 cycles of 30 s at 98°C, 30 s at 55°C, and 30 s at 72°C, followed by a final extension at 72°C for 5 min. Equal volumes (10uL) of individual BarSeq PCR's were pooled and purify with the DNA Clean and Concentrator kit (Zymo Reserch). The BarSeq library was sequenced in either HiSeq 2000 or HiSeq 2500 system (Illumina).

BarSeq Data analysis

The analysis was made according WETMORE (2015). Briefly, the reads were converted in tables showing the number of times each barcode was seen in each sample using the custom script (MultiCodes.pl). The analysis requires a perfect match of 9 nucleotides upstream of the barcode, otherwise the read is discarded. Knowing the table of barcodes, where they map in the genome and their counts in each sample was estimate the strains fitness

and gene fitness values with the custom R script FEBA.R. The strain fitness is the normalized log2 ratio of counts between the treatment sample and the reference "time-zero" sample. The gene fitness is the weighted average of the strain fitness. All the experiments were made with 2-5 replicates with independent gDNA extraction and independent PCR. Gene fitness ≥1 or ≤-1 was considered different. For the statistical analysis was used a moderated *t* statistic, genes with |t| of4 have highly significant phenotypes that are reproducible in biological replicate experiments (see WETMORE, 2015).

AUTHOR CONTRIBUTION

Adan Deutschbauer (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA) – Assisted the TnSeq library construction, sequencing and data analysis

Adan Arkin (Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA and Department of Bioengineering, University of California, Berkeley, California, USA) - Assisted the TnSeq library construction, sequencing and data analysis

Fernanda P do Amaral and Beverly Jose Agtuca (Division of Plant Science and Biochemistry, University of Missouri, USA) - Assisted the Sorghum Screening

Tomas P Pereira (Department of Science and Food Technology, Federal University of Santa Catarina, Brazil) - Assisted the Sorghum Screening

Fabio de Oliveira Pedrosa, Emanuel Maltempi de Souza and Rose Adele Monteiro (Department of Biochemistry and Molecular Biology, Federal University of Paraná, Brazil) – Orientation and Supervision

Gary Stacey (Division of Plant Science and Biochemistry, University of Missouri, USA) - Orientation and Supervision

REFERENCES

BALDANI, J.I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum* seropedicae gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. **International Journal of Systematic Bacteriology**. V. 36. P. 86-93. 1986.

BALDANI, V.L.D.; BALDANI, J.I.; OLIVARES, F.; DÖBEREINER, J. Identification and ecology of *Herbaspirillum* seropedicae and the closely related Pseudomonas rubrisubalbicans. **Symbiosis**, v. 13, p. 65-73, 1992.

BALDANI, V.L.D.; BALDANI, J.I.; DOBEREINER, J. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum* seropedicae and Burkholderia spp. **Biol. Fertil. Soils** 30:485-491, 2000.

BONATO, B., BATISTA, M. B. B., CAMILIOS-NETO, D., PANKIEVICZ, V. C., TADRA-SFEIR, M. Z., MONTEIRO, R. A., PEDROSA, F. O., SOUZA, E. M., CHUBATSU, L. S., WASSEM, R., RIGO, L. U. RNA-seq analyses reveal insights into the function of respiratory nitrate reductase of the diazotroph *Herbaspirillum* seropedica. **Environmental Microbiology** (2016) 18(8), 2677–2688

JAMES, E. K., OLIVARES, F. L., BALDANI, J. I., DÖBEREINER, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue 3Sorghum bicolor L. Moench **Journal of Experimental Botany** 48 (3), 785-798. 1997

KAVANAGHA, K. L., JÇRNVALLB, H., PERSSONC, B., OPPERMANNA, U. The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. **Cell.Mol.Life Sci.** 65 (2008) 3895–3906

MALIK, R., VIOLA, R. Structural characterization of tartrate dehydrogenase: a versatile enzyme catalyzing multiple reactions. **Biological Crystallography**. Volume 66, Part 6, 2010.

MORENO-VIVIÁN, C. et al. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. **Journal of Bacteriology**, v. 181, n. 21, p. 6573-84, Nov 1999

OLIVARES F. L.; JAMES E. K.; BALDANI J. I, DOBEREINER J. Infection of mottled stripe disease-susceptible and resistant sugar cane varieties by the endophytic diazotroph *Herbaspirillum*. **New Phytol**. vol.135, p. 723-737, 1997.

PEDROSA FO, MONTEIRO RA, WASSEM R, CRUZ LM et al (2011) Genome of *Herbaspirillum* seropedicae strain SmR1, a specialized diazotrophic endophyte of tropical grasses. **PLoS Genetics**. doi:10.1371/journal.pgen.1002064

RONCATO-MACCARI LDB, RAMOS HJO, PEDROSA FO, ALQUINI Y, CHUBATSU LS, YATES MG, RIGO LU, STEFFENS MBR, SOUZA EM. Root colonization, systemic spreading and contribution of *Herbaspirillum* seropedicae to growth of rice seedlings**. Symbiosis** 35:01–1. 2003.

RICHARDSON, A. E.; BAREA, J. M.; MCNEILL, A. M. Prigent-Combaret C. Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. Plant Soil.321:305–339. 2009.

SCHELL, M. A. Molecular biology of the LysR family of transcriptional regltlators. **Annu. Rev. Microbiol**. 47:597. 1993.

TIPTON PA, PEISACH J Characterization of the multiple catalytic activities of tartrate dehydrogenase**. Biochemistry**. Feb 20; 29(7):1749-56. 1990.

TIPTON, P. A. Tartrate dehydrogenase, an enzyme with multiple catalytic activities **Pept. Protein Lett.**7, 323–332. (2000).

WETMORE KM, PRICE MN, WATERS RJ, LAMSON JS, HE J, HOOVER CA, BLOW MJ, BRISTOW J, BUTLAND G, ARKIN AP, DEUTSCHBAUER A. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. **mBio** 6(3). 2015.

SUPPLEMENTARY TABLE 6-1- Exclusive sick genes found during the RB-TnSeq analysis for each Carbon source.

Locus Id	Gene annotation	Gene Fitness
Galactose (6)		
HSERO_RS09235	MarR family transcriptional regulator	-1.1255
HSERO_RS05155	2-dehydro-3-deoxy-6-phosphogalactonate aldolase	-5.7912
HSERO_RS05160	2-dehydro-3-deoxygalactonokinase	-3.1191
HSERO_RS13880	hypothetical protein	-1.3194
HSERO_RS21480	transcriptional regulator	-1.0337
HSERO_RS00410	phosphoheptose isomerase	-1.2491
Malato (16)		
HSERO_RS21740	methionine synthase	-2.0511
HSERO_RS13310	cobaltochelatase	-1.9284
HSERO_RS04230	LuxR family transcriptional regulator	-1.6871
HSERO_RS22025	aromatic amino acid aminotransferase	-1.1047
HSERO_RS07390	tartrate dehydrogenase	-3.0298
HSERO_RS15685	histidine kinase	-1.1712
HSERO_RS09770	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	-1.7015
HSERO_RS13280	precorrin-4 C11-methyltransferase	-2.2564
HSERO_RS10020	flippase	-1.0274
HSERO_RS07400	LysR family transcriptional regulator	-2.6228
HSERO_RS13265	G3E family GTPase	-1.4069
HSERO_RS13305	magnesium chelatase	-2.1542
HSERO_RS13240	cobalamin biosynthesis protein CobD	-2.1432
HSERO_RS13285	precorrin-2 C20-methyltransferase	-2.1454
HSERO_RS15680	chemotaxis protein CheY	-2.1302
HSERO_RS12920	trigger factor	-1.0839
Glucose (3)		
HSERO_RS15655	methylcitrate synthase	-1.2566
HSERO_RS15600	cytosol aminopeptidase	-1.5117
HSERO_RS19630	AMP nucleosidase	-1.1366
Lactato (28)		
HSERO_RS17105	DNA polymerase III subunit epsilon	-1.3079
HSERO_RS03045	HrcA family transcriptional regulator	-1.005
HSERO_RS09855	GTP pyrophosphokinase	-4.4218
HSERO_RS13960	hypothetical protein	-1.1069
HSERO_RS01510	membrane protein	-1.0669
HSERO_RS03520	aminoglycoside phosphotransferase	-1.3729
HSERO_RS16735	alcohol dehydrogenase	-1.7151

Continuation

HSERO_RS16745	gluconate 2-dehydrogenase	-1.901
HSERO_RS19105	FAD-binding protein	-1.274
HSERO_RS10560	membrane protein	-4.4767
HSERO_RS03440	biopolymer transporter ExbD	-1.8483
HSERO_RS05505	transcriptional regulator	-1.0497
HSERO_RS13650	high-affinity Fe2+/Pb2+ permease	-1.1885
HSERO_RS09230	2-dehydropantoate 2-reductase	-1.1808
HSERO_RS22285	GntR family transcriptional regulator	-1.2034
HSERO_RS19095	glycolate oxidase	-1.1134
HSERO_RS02910	IclR family transcriptional regulator	-1.1457
HSERO_RS16450	colicin V production protein	-1.0154
HSERO_RS02610	hypothetical protein	-1.0254
HSERO_RS21695	hypothetical protein	-1.0155
HSERO_RS16740	GMC family oxidoreductase	-2.1754
HSERO_RS16750	DeoR family transcriptional regulator	-1.0439
HSERO_RS21420	5-formyltetrahydrofolate cyclo-ligase	-1.0573
HSERO_RS21340	septum formation inhibitor protein	-1.0436
HSERO_RS02100	4-oxalocrotonate decarboxylase	-1.3023
HSERO_RS13655	hypothetical protein	-1.0685
HSERO_RS17900	3,4-dihydroxy-2-butanone 4-phosphate synthase	-2.6826
HSERO_RS13645	membrane protein	-1.1177
Manitol (12)		
HSERO_RS02210	sugar ABC transporter ATP-binding protein	-3.7159
HSERO_RS10230	flagellar hook-filament junction protein	-1.0784
HSERO_RS02195	sugar ABC transporter permease	-3.1723
HSERO_RS09455	NAD synthetase	-1.0297
HSERO_RS14930	aldolase	-1.1754
HSERO_RS07140	glutamine-synthetase adenylyltransferase	-1.095
HSERO_RS10355	flagellin	-1.0673
HSERO_RS02200	mannitol ABC transporter permease	-3.0981
HSERO_RS10225	flagellar hook protein FlgK	-1.1452
HSERO_RS02190	sugar ABC transporter substrate-binding protein	-3.392
HSERO_RS02220	xylulokinase	-2.9205
HSERO_RS10165	flagellar biosynthesis protein	-1.0162
Fucose (9)		
HSERO_RS13290	cobalt-precorrin-6A synthase	-1.0078
HSERO_RS17735	quinone oxidoreductase	-1.1184
HSERO_RS21425	5,10-methylenetetrahydrofolate reductase	-1.1725
HSERO_RS01735	ATPase AAA	-1.9699
HSERO_RS10320	flagellar brake protein YcgR	-1.1644

Continuation

HSERO_RS06350	3-oxoacyl-ACP reductase	-2.8338
HSERO_RS21765	cell division protein	-1.1266
HSERO_RS03265	transcription regulator protein	-1.0058
HSERO_RS09360	lysine transporter LysE	-2.5942
Manose (32)		
HSERO_RS03640	D-ribose transporter ATP-binding protein	-4.327
HSERO_RS14845	stationary phase survival protein SurE	-1.053
HSERO_RS02470	tRNA (guanine-N7)-methyltransferase	-1.5162
HSERO_RS12855	histidine kinase	-1.1469
HSERO_RS20600	transcription regulator protein	-1.6614
HSERO_RS03645	ribose ABC transporter permease	-2.9068
HSERO_RS10850	ABC transporter	-2.1638
HSERO_RS06530	muconolactone delta-isomerase	-1.3334
HSERO_RS22750	sugar ABC transporter ATP-binding protein	-1.1167
HSERO_RS22845	LysR family transcriptional regulator	-1.0432
HSERO_RS11265	GntR family transcriptional regulator	-1.0396
HSERO_RS20750	cytochrome C oxidase subunit I	-1.0173
HSERO_RS06145	NrdR family transcriptional regulator	-1.1196
HSERO_RS13965	single-stranded DNA exonuclease	-1.0172
HSERO_RS03650	N-acyl-D-glucosamine 2-epimerase	-5.4395
HSERO_RS03635	membrane protein	-1.0676
HSERO_RS20765	membrane protein	-1.0676
HSERO_RS00450	phosphoenolpyruvate-protein phosphotransferase	-2.4819
HSERO_RS02415	hemolysin D	-1.0161
HSERO_RS22755	ATPase	-1.0623
HSERO_RS06175	queuine tRNA-ribosyltransferase	-1.5851
HSERO_RS11495	LacI family transcription regulator	-2.9371
HSERO_RS11435	IclR family transcriptional regulator	-1.1822
HSERO_RS10340	flagellar biosynthesis protein FliS	-2.7685
HSERO_RS21470	hypothetical protein	-1.2202
HSERO_RS22450	aldose 1-epimerase	-1.1827
HSERO_RS21935	cytochrome C	-1.113
HSERO_RS03630	fructose transporter	-4.8826
HSERO_RS11500	ribokinase	-1.0655
HSERO_RS18685	hypothetical protein	-1.0081
HSERO_RS12105	RpiR family transcriptional regulator	-2.011
HSERO_RS00080	histidine kinase	-1.2471
Rhamnose (12)		
HSERO_RS22205	L-rhamnose mutarotase	-2.5024
HSERO_RS22210	branched-chain amino acid ABC transporter permease	-4.8552

		Conclusion
HSERO_RS22235	short-chain dehydrogenase	-5.9687
HSERO_RS22220	D-ribose transporter ATP-binding protein	-3.4763
HSERO_RS17120	hydroxyacylglutathione hydrolase	-1.5634
HSERO_RS01260	3-hydroxyacyl-CoA dehydrogenase	-1.7813
HSERO_RS22215	branched-chain amino acid ABC transporter permease	-4.1807
HSERO_RS22240	sugar isomerase	-4.6908
HSERO_RS19525	PTS sugar transporter subunit IIA	-1.4165
HSERO_RS22225	ABC transporter substrate-binding protein	-4.6929
HSERO_RS22230	DeoR faimly transcriptional regulator	-5.9429
HSERO_RS22200	L-fuculose kinase	-5.7305
Xylose (14)		
HSERO_RS02600	bile acid:sodium symporter	-1.9161
HSERO_RS19330	FAH family protein	-2.4133
HSERO_RS00735	2,5-dioxovalerate dehydrogenase	-2.5285
HSERO_RS16825	NADP-dependent malic enzyme oxidoreductase	-1.4
HSERO_RS19355	galactarate dehydratase	-1.9429
HSERO_RS19360	fumarylacetoacetate hydrolase	-4.3065
HSERO_RS19995	3-oxoadipate CoA-transferase subunit B	-1.6931
HSERO_RS19325	3-oxoadipate CoA-transferase subunit B	-1.6931
HSERO_RS05745	NADP-dependent malic enzyme oxidoreductase	-1.3156
HSERO_RS08610	anhydrase	-1.0859
HSERO_RS20265	hypothetical protein	-1.7814
HSERO_RS19335	ABC transporter permease	-1.7962
HSERO_RS15630	exodeoxyribonuclease III	-1.4313
HSERO_RS18365	hypothetical protein	-1.0266

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-2 – Exclusive sick genes found during the RB-TnSeq analysis for Osmotic Stress (NaCl 1%).

Locus Id	Gene annotation	Gene Fitness
NaCl (39)		
HSERO_RS00975	alpha,alpha-trehalose-phosphate synthase	-5,4249
HSERO_RS13105	C4-dicarboxylate ABC transporter	-2,9610
HSERO_RS03595	cation transporter	-2,1488
HSERO_RS00970	glucoamylase	-2,0929
HSERO_RS04235	C4-dicarboxylate ABC transporter	-2,0324
HSERO_RS19630	AMP nucleosidase	-1,9320
HSERO_RS21450	lipid A biosynthesis acyltransferase	-1,7192
HSERO_RS16745	gluconate 2-dehydrogenase	-1,6233
HSERO_RS15545	membrane protein	-1,6222
HSERO_RS13655	hypothetical protein	-1,6099
HSERO_RS20625	phenylacetate-CoA oxygenase	-1,5670
HSERO_RS18455	phosphoglycolate phosphatase	-1,5323
HSERO_RS09660	ribonuclease III	-1,5002
HSERO_RS09145	DNA topoisomerase IV subunit A	-1,4899
HSERO_RS16075	coproporphyrinogen III oxidase	-1,4435
HSERO_RS23945	cobyric acid synthase CobQ	-1,4244
HSERO_RS16835	phosphatidylglycerophosphatase	-1,4149
HSERO_RS10400	inorganic phosphate transporter	-1,4056
HSERO_RS16740	GMC family oxidoreductase	-1,3789
HSERO_RS08605	Hsp33 chaperonin	-1,3343
HSERO_RS13650	high-affinity Fe2+/Pb2+ permease	-1,3266
HSERO_RS18365	hypothetical protein	-1,3106
HSERO_RS09795	carbon starvation protein A	-1,2724
HSERO_RS16660	NAD-dependent dehydratase	-1,2515
HSERO_RS09800	hypothetical protein	-1,2343
HSERO_RS16735	alcohol dehydrogenase	-1,2303
HSERO_RS21445	lipid A biosynthesis acyltransferase	-1,1763
HSERO_RS15040	poly(R)-hydroxyalkanoic acid synthase	-1,1403
HSERO_RS21550	phosphoenolpyruvate carboxykinase	-1,1266
HSERO_RS01260	3-hydroxyacyl-CoA dehydrogenase	-1,1098
HSERO_RS09155	DNA topoisomerase IV subunit B	-1,1034
HSERO_RS10865	PII uridylyl-transferase	-1,0981
HSERO_RS09785	LuxR family transcriptional regulator	-1,0942
HSERO_RS20875	hypothetical protein	-1,0802
HSERO_RS08640	hypothetical protein	-1,0670
HSERO_RS10960	PEP synthetase regulatory protein	-1,0502
HSERO_RS08405	membrane protein	-1,0404

		Conclusion
HSERO_RS13965	single-stranded DNA exonuclease	-1,0350
HSERO_RS13645	membrane protein	-1,0087

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-3 – Exclusive sick genes found during the RB-TnSeq analysis in the indicated Nitrogen Source under High aeration (HA).

Gene Id	**Gene anotattion**	***Gene fitness***
Glutamate High Aeration (17)		
HSERO_RS19090	C4-dicarboxylate ABC transporter	-2,6669
HSERO_RS05745	NADP-dependent malic enzyme oxidoreductase	-1,7526
HSERO_RS02760	hypothetical protein	-1,6335
HSERO_RS02255	decarboxylase	-1,5180
HSERO_RS19240	arginine ABC transporter ATP-binding protein	-1,4583
HSERO_RS01315	transcriptional regulator	-1,4116
HSERO_RS19250	glutamate ABC transporter permease	-1,3090
HSERO_RS16860	hypothetical protein	-1,3072
HSERO_RS03915	hypothetical protein	-1,2625
HSERO_RS19255	ABC transporter	-1,2254
HSERO_RS18875	ATPase	-1,2183
HSERO_RS14840	protein-L-isoaspartate O-methyltransferase	-1,1952
HSERO_RS02205	phosphatase	-1,1372
HSERO_RS09235	MarR family transcriptional regulator	-1,1259
HSERO_RS08975	membrane protein	-1,0418
HSERO_RS20250	large conductance mechanosensitive channel protein MscL	-1,0382
HSERO_RS00765	cytochrome C	-1,0141
Nitrate High Aeration (44)		
HSERO_RS14540	nitrite reductase	-4,0766
HSERO_RS16660	NAD-dependent dehydratase	-3,8100
HSERO_RS14535	nitrite reductase	-3,6489
HSERO_RS22335	nitrate response regulator protein	-3,1403
HSERO_RS02140	thiamine-phosphate synthase	-2,8682
HSERO_RS14530	MFS transporter	-2,7517
HSERO_RS09255	molybdenum cofactor biosynthesis protein MogA	-2,6027
HSERO_RS10495	molybdenum cofactor biosynthesis protein MoaE	-2,5481
HSERO_RS05910	phosphoribosylglycinamide formyltransferase	-2,5478
HSERO_RS00835	molybdopterin biosynthesis protein MoeB	-2,5021
HSERO_RS09400	molybdopterin-guanine dinucleotide biosynthesis protein	-2,3037
HSERO_RS14550	nitrate reductase	-2,2916
HSERO_RS02130	thiazole synthase	-2,2853
HSERO_RS02120	phosphomethylpyrimidine synthase ThiC	-2,2001
HSERO_RS02135	phosphomethylpyrimidine kinase	-2,1153
HSERO_RS14545	nitrite reductase	-2,0845
HSERO_RS11085	L-aspartate oxidase	-2,0649
HSERO_RS02125	thiamine biosynthesis protein ThiS	-2,0609

		Conclusion
HSERO_RS03300	molybdenum cofactor biosynthesis protein MoaC	-1,9753
HSERO_RS09395	molybdenum cofactor biosynthesis protein MoaA	-1,8130
HSERO_RS13655	hypothetical protein	-1,7019
HSERO_RS00965	trehalose-phosphatase	-1,6614
HSERO_RS13660	membrane protein	-1,6591
HSERO_RS09405	molybdenum cofactor biosynthesis protein MoeA	-1,6551
HSERO_RS21580	membrane protein	-1,5409
HSERO_RS13645	membrane protein	-1,4242
HSERO_RS00425	ammonia channel protein	-1,3838
HSERO_RS22980	membrane protein	-1,3457
HSERO_RS04235	C4-dicarboxylate ABC transporter	-1,2890
HSERO_RS05500	transaldolase	-1,2861
HSERO_RS00430	glutamate--cysteine ligase	-1,2797
HSERO_RS06650	MarR family transcriptional regulator	-1,2660
HSERO_RS21740	methionine synthase	-1,1811
HSERO_RS21375	N5-glutamine adenosylmeth-dependent methyltransferase	1,1794
HSERO_RS18455	phosphoglycolate phosphatase	-1,1512
HSERO_RS10395	phosphate transport regulator	-1,1440
HSERO_RS13310	cobaltochelatase	-1,1361
HSERO_RS13650	high-affinity Fe2+/Pb2+ permease	-1,1242
HSERO_RS21935	cytochrome C	-1,0820
HSERO_RS10815	chromosome segregation protein SMC	-1,0529
HSERO_RS07275	D-amino acid dehydrogenase	-1,0421
HSERO_RS00815	sulfurtransferase	-1,0419
HSERO_RS23945	cobyric acid synthase CobQ	-1,0164
HSERO_RS19815	phytanoyl-CoA dioxygenase	-1,0011
NH_4Cl High Aeration (11)		
HSERO_RS20985	serine acetyltransferase	-1,6338
HSERO_RS20265	hypothetical protein	-1,4026
HSERO_RS20240	cytochrome B	-1,2672
HSERO_RS00220	hypothetical protein	-1,2557
HSERO_RS13280	precorrin-4 C11-methyltransferase	-1,1716
HSERO_RS21475	recombinase XerC	-1,1432
HSERO_RS12855	histidine kinase	-1,0584
HSERO_RS20820	ABC transporter ATP-binding protein	-1,0495
HSERO_RS06420	hypothetical protein	-1,0215
HSERO_RS09660	ribonuclease III	-1,0199
HSERO_RS10805	succinyldiaminopimelate aminotransferase	-1,0058

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-4 – Exclusive sick genes found during the RB-TnSeq analysis in the indicated Nitrogen Source under Low aeration (LA).

Gene Id	Gene anotattion	*Gene fitness*
Glutamate LA (14)		
HSERO_RS19240	arginine transporter atp-binding subunit	-1,71182
HSERO_RS19250	glutamate abc transporter permease	-1,13079
HSERO_RS20985	transferase	-2,65692
HSERO_RS19090	c4-dicarboxylate abc transporter	-1,92797
HSERO_RS21945	hypothetical protein	-1,09367
HSERO_RS08405	membrane protein	-1,25719
HSERO_RS20135	type VI secretion protein	-1,04383
HSERO_RS19245	glutamate aspartate transporter permease	-1,25583
HSERO_RS05745	malic enzyme	-1,47225
HSERO_RS22350	urea carboxylase	-1,66511
HSERO_RS19630	amp nucleosidase	-1,8203
HSERO_RS21340	cell division protein	-1,00985
HSERO_RS02760	pf05974 domain protein	-1,00759
HSERO_RS07970	formyltetrahydrofolate deformylase	-2,40039
Nitrate LA (53)		
HSERO_RS09660	ribonuclease III	-1,39878
HSERO_RS21740	methionine synthase	-1,92218
HSERO_RS13310	cobalt chelatase	-1,58977
HSERO_RS08075	Ferredoxin	-1,18215
HSERO_RS11085	l-aspartate oxidase	-2,6751
HSERO_RS00835	molybdopterin biosynthesis protein	-2,08425
HSERO_RS08515	diguanylate cyclase	-1,24344
HSERO_RS21475	tyrosine recombinase	-1,23302
HSERO_RS04230	histidine kinase	-2,8049
HSERO_RS13650	FTR1 family iron permease	-1,12696
HSERO_RS09235	family transcriptional regulator	-1,18188
HSERO_RS02130	thiazole synthase	-2,05458
HSERO_RS00425	ammonia channel protein	-1,36033
HSERO_RS05500	Transaldolase	-1,00163
HSERO_RS15545	membrane protein	-1,32286
HSERO_RS13105	c4-dicarboxylate transporter	-1,07232
HSERO_RS06530	muconolactone delta-isomerase	-1,54802
HSERO_RS13660	membrane protein	-1,48078
HSERO_RS14950	chemotaxis protein	-1,02801
HSERO_RS22990	family transcriptional regulator	-1,73171
HSERO_RS13265	cobalamin biosynthesis protein	-1,62348

Conclusion

HSERO_RS13305	magnesium chelatase	-1,75717
HSERO_RS14540	nitrite reductase	-4,38844
HSERO_RS02140	thiamine-phosphate pyrophosphorylase	-1,48851
HSERO_RS09770	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	-2,2026
HSERO_RS16075	coproporphyrinogen III oxidase	-1,65664
HSERO_RS02160	DNA helicase	-1,26028
HSERO_RS13315	Ferredoxin	-2,11001
HSERO_RS16840	competence damage-inducible protein a	-1,10931
HSERO_RS04235	C4-dicarboxylate ABC transporter	-1,90789
HSERO_RS00430	glutamate--cysteine ligase	-1,48432
HSERO_RS13240	cobalamin biosynthesis protein	-1,5356
HSERO_RS10815	chromosome segregation protein smc	-1,49957
HSERO_RS10395	phosphate transport regulator	-1,11785
HSERO_RS02125	thiamine biosynthesis protein	-1,77156
HSERO_RS05620	dehydratase	-1,14243
HSERO_RS07600	cobalamin biosynthesis protein	-1,03665
HSERO_RS15070	pseudouridine synthase	-1,18641
HSERO_RS13280	precorrin-4 c11-methyltransferase	-1,78175
HSERO_RS17110	ribonuclease h	-1,50408
HSERO_RS08655	acetolactate synthase	-1,2419
HSERO_RS14530	MFS transporter	-2,94955
HSERO_RS21580	membrane protein	-2,0262
HSERO_RS09220	hypothetical protein	-1,53822
HSERO_RS06650	marR family transcriptional regulator	-1,4609
HSERO_RS02135	phosphomethylpyrimidine kinase	-1,69068
HSERO_RS02120	thiamine biosynthesis protein	-1,90293
HSERO_RS14535	nitrite reductase	-4,12903
HSERO_RS22335	antar domain-containing protein	-3,04148
HSERO_RS13285	precorrin-2 c20-methyltransferase	-1,5809
HSERO_RS22980	membrane protein	-1,70432
HSERO_RS13655	periplasmic lipoprotein involved in iron transport	-1,27957
HSERO_RS16660	nad-dependent dehydratase	-3,57932
NH_4Cl (3)		
HSERO_RS02215	D-arabinitol 4-dehydrogenase	-1,2378
HSERO_RS00965	trehalose phosphatase	-1,96156
HSERO_RS19995	3-oxoadipate -transferase	-1,00511

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-5 – Exclusive sick genes found during the RB-TnSeq analysis in Glutamate as Nitrogen Source under High (HA) and Low aeration (LA).

Gene Id	**Gene anotattion**	***Gene fitness***
Glutamate (HA)		
HSERO_RS09770	nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	-1,128531856
HSERO_RS00765	cytochrome C	-1,014092963
HSERO_RS02160	DNA-dependent helicase	-1,307982396
HSERO_RS13315	ferredoxin	-1,617198702
HSERO_RS01315	transcriptional regulator	-1,411631234
HSERO_RS10455	aminotransferase	-1,012314109
HSERO_RS13240	cobalamin biosynthesis protein CobD	-1,387007202
HSERO_RS13260	cobinamide adenolsyltransferase	-1,093223257
HSERO_RS03915	hypothetical protein	-1,262537472
HSERO_RS19255	ABC transporter	-1,225429054
HSERO_RS15070	ribosomal large subunit pseudouridine synthase D	-1,141253136
HSERO_RS17110	ribonuclease	-1,281387079
HSERO_RS14840	protein-L-isoaspartate O-methyltransferase	-1,195206916
HSERO_RS02255	decarboxylase	-1,518028298
HSERO_RS16860	hypothetical protein	-1,307187887
HSERO_RS08975	membrane protein	-1,041800964
HSERO_RS09235	MarR family transcriptional regulator	-1,125927582
HSERO_RS02205	phosphatase	-1,137199846
HSERO_RS02090	Holliday junction DNA helicase RuvA	-2,052378914
HSERO_RS20250	large conductance mechanosensitive channel protein MscL	-1,038199972
HSERO_RS11065	quinolinate synthetase	-1,101721849
HSERO_RS04265	tyrosine recombinase XerD	-1,248469694
HSERO_RS18875	ATPase	-1,218286147
HSERO_RS13305	magnesium chelatase	-1,257004801
Glutamate (LA)		
HSERO_RS19245	glutamate/aspartate transporter permease GltK	-1,255830791
HSERO_RS13255	cobyrinic acid a,c-diamide synthase	-1,017142053
HSERO_RS22350	urea carboxylase	-1,665108736
HSERO_RS19630	AMP nucleosidase	-1,820296323
HSERO_RS21340	septum formation inhibitor protein	-1,009854896
HSERO_RS20985	serine acetyltransferase	-2,656916543
HSERO_RS07970	formyltetrahydrofolate deformylase	-2,400388391
HSERO_RS21945	hypothetical protein	-1,093671939
HSERO_RS08640	hypothetical protein	-1,069477692
HSERO_RS20135	type VI secretion protein	-1,043829947

HSERO_RS13270	bifunctional precorrin-3 reductase oxidoreductase		methyltransferase/precorin-6x	-1,024618184

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-6 – Exclusive sick genes found during the RB-TnSeq analysis in Nitrate as Nitrogen Source under High (HA) and Low aeration (LA).

Gene Id	**Gene anotattion**	***Gene fitness***
Nitrate (HA)		
HSERO_RS14545	nitrite reductase	-2,084478543
HSERO_RS03300	molybdenum cofactor biosynthesis protein MoaC	-1,97534989
HSERO_RS19815	phytanoyl-CoA dioxygenase	-1,001109046
HSERO_RS23945	cobyric acid synthase CobQ	-1,016434612
HSERO_RS10495	molybdenum cofactor biosynthesis protein MoaE	-2,548082624
HSERO_RS21375	N5-glutamine S-adenosyl-L-methionine-dependent methyltransferase	-1,179388087
HSERO_RS08405	membrane protein	-1,206524383
HSERO_RS02090	Holliday junction DNA helicase RuvA	-1,691347256
HSERO_RS09255	molybdenum cofactor biosynthesis protein MogA	-2,602699474
HSERO_RS01260	3-hydroxyacyl-CoA dehydrogenase	-1,485016653
HSERO_RS03085	3-methyl-2-oxobutanoate hydroxymethyltransferase	-1,330456571
HSERO_RS07275	D-amino acid dehydrogenase	-1,042122554
HSERO_RS00965	trehalose-phosphatase	-1,661392342
HSERO_RS14550	nitrate reductase	-2,291567532
HSERO_RS21935	cytochrome C	-1,082032723
HSERO_RS09400	molybdopterin-guanine dinucleotide biosynthesis protein A	-2,303704069
HSERO_RS09395	molybdenum cofactor biosynthesis protein MoaA	-1,812991262
HSERO_RS13965	single-stranded DNA exonuclease	-1,415837648
HSERO_RS00815	sulfurtransferase	-1,041914986
HSERO_RS18455	phosphoglycolate phosphatase	-1,151203277
HSERO_RS09405	molybdenum cofactor biosynthesis protein MoeA	-1,655081604
HSERO_RS13645	membrane protein	-1,424221611
Nitrate (LA)		
HSERO_RS09660	ribonuclease III	-1,398776064
HSERO_RS08075	ferredoxin	-1,182152791
HSERO_RS09775	cobalamin synthase	-1,789439409
HSERO_RS08515	diguanylate cyclase	-1,243441143
HSERO_RS21475	recombinase XerC	-1,233021061
HSERO_RS09235	MarR family transcriptional regulator	-1,181875343
HSERO_RS15545	membrane protein	-1,322858676
HSERO_RS13105	C4-dicarboxylate ABC transporter	-1,07231674
HSERO_RS06530	muconolactone delta-isomerase	-1,548018828
HSERO_RS14950	purine-binding chemotaxis protein	-1,028007467
HSERO_RS08115	acyl-CoA-binding protein	-1,019371602
HSERO_RS13305	magnesium chelatase	-1,757174084

Conclusion

HSERO_RS16075	coproporphyrinogen III oxidase	-1,656636812
HSERO_RS13315	ferredoxin	-2,110005268
HSERO_RS16840	competence damage-inducible protein A	-1,109306816
HSERO_RS05620	dehydratase	-1,142426311
HSERO_RS07600	threonine-phosphate decarboxylase	-1,036646462
HSERO_RS15070	ribosomal large subunit pseudouridine synthase D	-1,186408535
HSERO_RS13280	precorrin-4 C11-methyltransferase	-1,781751706
HSERO_RS09220	hypothetical protein	-1,538219213
HSERO_RS13285	precorrin-2 C20-methyltransferase	-1,580899339

SOURCE: The Author (2017).

SUPPLEMENTARY TABLE 6-7– Exclusive sick genes found during the RB-TnSeq analysis in Ammonium Chloride as Nitrogen Source under High (HA) and Low aeration (LA).

Gene Id	Gene anotattion	*Gene fitness*
Ammonium Chloride (HA)		
HSERO_RS09660	ribonuclease III	-1,019922736
HSERO_RS02160	DNA-dependent helicase	-1,251011496
HSERO_RS20820	ABC transporter ATP-binding protein	-1,049528882
HSERO_RS13315	ferredoxin	-1,24533315
HSERO_RS13320	membrane protein	-1,075320808
HSERO_RS13255	cobyrinic acid a,c-diamide synthase	-1,030288871
HSERO_RS05515	glucose-6-phosphate isomerase	-2,004073508
HSERO_RS21475	recombinase XerC	-1,143199121
HSERO_RS04230	LuxR family transcriptional regulator	-1,139289984
HSERO_RS20265	hypothetical protein	-1,402625168
HSERO_RS13280	precorrin-4 C11-methyltransferase	-1,171600104
HSERO_RS15070	ribosomal large subunit pseudouridine synthase D	-1,216391398
HSERO_RS17110	ribonuclease	-1,36621719
HSERO_RS13965	single-stranded DNA exonuclease	-1,307214542
HSERO_RS20985	serine acetyltransferase	-1,633821
HSERO_RS08655	acetolactate synthase	-2,022448787
HSERO_RS12855	histidine kinase	-1,058424299
HSERO_RS08405	membrane protein	-1,278308122
HSERO_RS20240	cytochrome B	-1,267231362
HSERO_RS08640	hypothetical protein	-1,034956525
HSERO_RS06420	hypothetical protein	-1,021488798
HSERO_RS10805	succinyldiaminopimelate aminotransferase	-1,005770545
HSERO_RS01260	3-hydroxyacyl-CoA dehydrogenase	-1,019991148
HSERO_RS22990	transcriptional regulator	-1,011550752
HSERO_RS03085	3-methyl-2-oxobutanoate hydroxymethyltransferase	-1,162228355
HSERO_RS13265	G3E family GTPase	-1,177382237
HSERO_RS00220	hypothetical protein	-1,255711025
HSERO_RS04265	tyrosine recombinase XerD	-1,031913833
HSERO_RS13305	magnesium chelatase	-1,203910039
Ammonium Chloride (LA)		
HSERO_RS00965	trehalose-phosphatase	-1,961562711
HSERO_RS19995	3-oxoadipate CoA-transferase subunit B	-1,005110501
HSERO_RS05910	phosphoribosylglycinamide formyltransferase	-1,387442582
HSERO_RS13260	cobinamide adenolsyltransferase	-1,01554163
HSERO_RS02215	D-arabinitol 4-dehydrogenase	-1,237797897
HSERO_RS13295	membrane protein	-1,137886067

		Conclusion
HSERO_RS08115	acyl-CoA-binding protein	-1,274790831

SOURCE: The Author (2017).

CAPÍTULO VII – CONCLUSÕES

Este trabalho teve como objetivo identificar fatores envolvidos durante a interação planta bactéria entre bactérias do gênero *Herbaspirillum* e gramíneas. A identidicação prévia de um cluster de biossíntese se celulose presente em *H. rubrisubalbicans* e ausente em *H. seropedicae* e o estudo da função deste cluster demostrou que a produção deste EPS é importante para a formação de biofilme bacteriano e está relacionado com a colonização epifítica de *H rubrisubalbicans* em gramíneas.

O estudo da interação patogênica entre *H rubrisubalbicans* e sorgo demostrou diversos fatores envolvidos com a colonização e desenvolvimento da doença da estria vermelha. O envolvimento do sistema imune vegetal parece ser essencial para o controle da colonização bacteriana nos tecidos internos e desenvolvimento dos sintomas da doença. Análises de QTL e RNASeq apontaram genes envolvidos com a modulação do sistema imune e que parecem estar diretamente relacionados com a infecção dos tecidos vegetais por *H. rubrisubalbicans* M1. Proteínas envolvidas nas vias de sinalização por íons cálcio, receptores envolvidos no reconhecimento de quitina e peptiodeoglicanos da parede celular bacteriana (CERK1), e proteínas envolvidas com vias moduladas pelo Sistema de Secreção do tipo III bacteriano, como RIN4 e HSP90 também foram identificadas durante a associação entre *H. rubrisubalbicans* e Sorgo.

Ensaios de TnSeq com o objetivo de analisar a interação benéfica entre *Herbaspirillum seropedicae* SmR1 e *Setaria viridis* foram realizados e indicaram diversos fatores relacionados a esta interação, como a importância do aparato flagelar, do sistema de quimiotaxia, da organização da parede celular e do metabolismo de PHB. O mesmo ensaio de TnSeq foi realizado em culturas puras, onde houve variação das fontes de Carbono, nitrogênio e estresse osmótico. Este ensaio possibilitou a identificação de diversos transportadores e reguladores transcricionais, bem como uma grande quantidade de proteínas hipotéticas, demostrando um grande potencial desta técnica para anotação de genes com função desconhecida.

CAPÍTULO VIII – REFERENCIAS

AUSMEES, N., JONSSON, H., HOGLUND, S., LJUNGGREN, H. AND LINDBERG, M. Structural and putative regulatory genes involved in cellulose synthesis in Rhizobium leguminosarum bv. trifolii. Microbiology 145:1253-1262. 1999.

BALDANI, J.I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum* seropedicae gen. nov., sp. nov., a root-associated nitrogen-fixing bacterium. **International Journal of Systematic Bacteriology**. V. 36. P. 86-93. 1986.

BALDANI, V.L.D.; BALDANI, J.I.; OLIVARES, F.; DÖBEREINER, J. Identification and ecology of *Herbaspirillum* seropedicae and the closely related Pseudomonas rubrisubalbicans. **Symbiosis**, v. 13, p. 65-73, 1992.

BALDANI B., POT G., KIRCHHOF E., FALSEN V. L. D., BALDANI F. L., OLIVARES, B. HOSTE, K. KERSTERS, A. HARTMANN, M. GILLIS, J. DÖBEREINER. Emended Description of *Herbaspirillum*; Inclusion of [Pseudomonas] rubrisubalbicans, a Mild Plant Pathogen, as *Herbaspirillum* rubrisubalbicans comb. nov.; and Classification of a Group of Clinical Isolates (EF Group 1) as *Herbaspirillum* Species 3. **Int J Syst Bacteriol,** v. 46, p.802- 810, 1996.

BALDANI, V.L.D.; BALDANI, J.I.; DOBEREINER, J. Inoculation of rice plants with theendophytic diazotrophs *Herbaspirillum* seropedicae and Burkholderia spp. **Biol. Fertil. Soils** 30:485-491, 2000.

BALSANELLI, E., TADRA-SFEIR, M. Z., FAORO, H., PANKIEVICZ, V., BAURA, V. A., PEDROSA, F. O., SOUZA, E. M. DIXON, R. & MONTEIRO, R. A. Molecular adaptations of *Herbaspirillum* seropedicae during colonization of the maize rhizosphere. **Environmental microbiology**. (2015).

BERG G, EBERL L, HARTMANN A (2005) The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ Microbiol 7:1673–1685

BODDEY RM, DE OLIVEIRA OC, URQUIAGA S, REIS VM, OLIVARES FL, BALDANI VLD, DÖBEREINER J (1995) Biological nitrogen fixation associated with sugar cane and rice: contributions and prospects for improvement. **Plant Soil** 174:195–209

BOYER JS, Plant productivity and environment. **Science** 218:443–448 (1982).

CRUZ LM, SOUZA EM, WEBER OB, BALDANI JI, DÖBEREINER J, PEDROSA FO. 16S Ribosomal DNA characterization of nitrogen-fixing Bacteria isolated from banana (Musa spp.) and pineapple (Ananas comosus (L.) Merril). **Appl Environ Microbiol** 67:2375–2379. 2001

DONG, T.G., HO, B.T., YODER-HIMES, D.R., AND MEKALANOS, J.J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in Vibrio cholerae. **Proc Natl Acad Sci USA** 110: 2623–2628.

ELBELTAGY A, NISHIOKA K, SATO T Endophytic colonization and in plant nitrogen fixation by a *Herbaspirillum* sp. isolated from wild rice species. **Appl Environ Microbiol** 67:5285–5293 (2001)

FUQUA, THOMAS e CLAY. Biofilm Formation by Plant-Associated Bacteria. **Annu. Rev. Microbiol.** 61:401–22. 2007

GALLI, F.; CARVALHO, P.C.T.; TOKESHI, H.; BALMER, F.; KIMATI, H.; CARDOSO, C.O.; SALGASO, C.L.; KRUGNER, T.L.; CARDOSO, E. J. B. N.; BERGAMIN, F. A. Manual de fitopatologia: doenças de plantas cultivadas. **Agronômica Ceres**, São Paulo (1980)

GYANESHWAR P, JAMES EK, REDDY PM, LADHA JK (2002) *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. **New Phytol** 154:131–145

HALE, C. N., AND J. P. WILKE. 1972. A comparative study of Pseudomonas species pathogenic to sorghum**. N. Z. J. Agric. Re**s. 15:448–456.

JAMES, E. K. Nitrogen fixation in endophytic and associative symbiosis. **Field Crops Research.** Volume 65, Issues 2–3, March, Pages 197–209. 2000

JAMES EK, GYANESHWAR P, MATHAN N Infection and colonization of rice seedlings by the plant growthpromotion bacterium *Herbaspirillum* seropedicae Z67. **Mol Plant Microbe Interact** 15:894–906 (2002)

JAHN, A. et al. Composition of Pseudomonas putida biofilms: accumulation of protein in the biofilm matrix. **Biofouling** 14, 49–57. 2000.

JAMES, E. K.; OLIVARES, F. L.; BALDANI, J. I.; DÖBEREINER, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of Sorghum bicolor L. Moench. **Journal of Experimental Botany**. v. 48. n. 308. p. 785-797. 1997.

KANG, K. S; COTTRELL, I. W. Em Polysaccharides in Microbial Technology, Peppler, H. J.; Perlman, D., eds.; **Academic Press: New York**, 1979, p. 417

MATTHYSSE, A.G., HOLMES, K.V. AND GURLITZ, R.H. Elaboration of cellulose fibrils by Agrobacterium tumefaciens during attachment to carrot cells. **J Bacteriol** 145:583-595. 1981.

MCBEE, GG, MILLER FR, DOMINY RE AND MONK RL, Quality of sorghum biomass for methanogenesis. **Energy from Biomass and Waste X**, ed. by Klass DL. Elsevier, London. pp. 251–260 (1987)

MONTEIRO, R. A., BALSANELLI, E.; TULESKI, T. R.; FAORO, H.; CRUZ, L. M.; WASSEM, R.; BAURA, V. A,; TADRA-SFEIR, M. Z.; WEISS, V.; DAROCHA, W. D.; MULLER-SANTOS, M.; CHUBATSU, L. S.; HUERGO, L. F.; PEDROSA, F. O. AND SOUZA, E. M. Genomic comparison of the endophyte *Herbaspirillum* seropedicae SmR1 and the phytopathogen *Herbaspirillum* rubrisubalbicans M1 by Suppressive Subtractive Hybridization and partial genome sequencing. **FEMS Microbiology Ecology,** 2012.

OLIVARES F. L.; JAMES E. K.; BALDANI J. I, DOBEREINER J. Infection of mottled stripe disease-susceptible and resistant sugar cane varieties by the endophytic diazotroph *Herbaspirillum*. **New Phytol**. vol.135, p. 723-737, 1997.

OPIJNEN, T. V. AND CAMILLI, A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. **Nat Rev Microbiol** . 2013 July ; 11(7).

PANKIEVICZ, V. AMARAL,,F. P., SANTOS, K., AGTUCA, B., XU, Y., SCHUELLER, M. J., ARISI, A. C., STEFFENS, M. B., SOUZA, E. M., PEDROSA, F. O., STACEY, G., FERRIERI, R. Robust biological nitrogen fixation in a model grass–bacterial association. **The Plant Journal** 81, 907–919. 2015

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., SOUZA E. M. RNA-seq transcriptional profiling of *Herbaspirillum* seropedicae colonizing wheat (Triticum aestivum) roots. **Plant Mol Biol**, 2016.

PIMENTEL, J.P., OLIVARES, F., PITARD, R. M., URQUIAGA, S., AKIBA, F., DOBEREINER, J. Dinitrogen fixation and infection of grass leaves by Pseudomonas rubrisubalbicans and *Herbaspirillum* seropedicae **Plant and Soil**, 1991

RONCATO-MACCARI, L. D. B., RAMOS, H. J.O., PEDROSA, F. O., ALQUINI, Y., CHUBATSU, L. S., YATES, M.G., RIGO, L.U.,STEFFENS, M.B., SOUZA, E. M. Endophytic *Herbaspirillum* seropedicae expresses nif genes in gramineous plants. **FEMS Microbiology ecology**, 45(37-47), 2003 62

RICHARDSON, A. E.; BAREA, J. M.; MCNEILL, A. M. Prigent-Combaret C. Acquisition of phosphorus and nitrogen in the rhizosphere and plant growth promotion by microorganisms. Plant Soil.321:305–339. 2009.

SOLANO, C.; GARCÍA, B.; VALLE, J., BERASAIN, C.; GHIGO, J.; GAMAZO, C.; LASA, I.; Genetic analusis of Salmonella enteritidis biofilm formation: critical role of cellulose. **Molecular microbiology.** 43(3), 793-808; 2002.

ROONEY, W. L., BLUMENTHAL, J., BEAN, B. AND MULLET, J. E. (2007), Designing sorghum as a dedicated bioenergy feedstock. **Biofuels, Bioprod. Bioref.,** 1: 147–157.

STURZ AV, CHRISTIE BR, NOWAK J (2000) Bacterial endophytes: potential role in developing sustainable systems of crop production. **CRC Crit Rev Plant Sci** 19:1–30

TAN, Z. Q., MEN, R., ZHANG, R. Y., HUANG, Z. First Report of *Herbaspirillum rubrisubalbicans* Causing Mottled Stripe Disease on Sugarcane in China. **Plant disease - Disease Notes**. March 2010, Volume 94, Number 3

WALKER T.S.; BAIS, H.P.; DEZIEL, E.; SCHWEIZER, H.P.; RAHME, L.G.; Pseudomonas aeruginosa-plant root interactions. Pathogenicity, biofilm formation, and root exudation. **Plant Physiology**; 134:320–31. 2004

WATNICK, P. And KOLTER, R. MINIREVIEW - Biofilm, City of Microbes. **Journal of bacteriology**., p. 2675–2679 Vol. 182, No. 10. 2000.

WETMORE KM, PRICE MN, WATERS RJ, LAMSON JS, HE J, HOOVER CA, BLOW MJ, BRISTOW J, BUTLAND G, ARKIN AP, DEUTSCHBAUER A. Rapid quantification of mutant fitness in diverse bacteria by sequencing randomly bar-coded transposons. **mBio** 6(3), 2015.

WIEDENFELD, R.P. Nutrient requirements and use efficiency by sweet sorghum. **Energy in Agriculture**, 3 (1984) 49--59 49