Universidade Federal do Rio Grande do Sul
Centro de Biotecnologia

Estudo da bactéria promotora de crescimento vegetal *Azospirillum amazonense*: aspectos genômicos e ferramentas genéticas específicas

Fernando Hayashi Sant'Anna

Porto Alegre, 2011

Universidade Federal do Rio Grande do Sul
Centro de Biotecnologia

Estudo da bactéria promotora de crescimento vegetal *Azospirillum amazonense*: aspectos genômicos e ferramentas genéticas específicas

Fernando Hayashi Sant'Anna

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

Orientadora: Profa. Dra. Irene Silveira Schrank

Porto Alegre, 2011

Este trabalho foi realizado no Laboratório de Microrganismos Diazotróficos, sendo financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Dedico esta tese aos meus pais, Paulo e Luciana, e ao meu irmão, Eduardo.

“Acredito muito na sorte; percebo que, quanto mais trabalho, mais a sorte me sorri.”

Thomas Jefferson

AGRADECIMENTOS

"Olhe só para mim, sou um pós-graduando. Tenho 30 anos e no ano passado ganhei 600 doláres." "Bart, não zombe dos pós-graduandos, eles apenas escolheram uma vida terrível." O diálogo ocorrido entre Marge e seu filho Bart em um episódio de "Os Simpsons" tem seu fundo de verdade. Vida de pós-graduando não é fácil. De doutorando então... Por ter vivido isso e por conhecer pessoas na mesma situação, reconheço: *pode* ser muito frustrante. Muito estudo, muito trabalho, poucos resultados, pouco retorno. Contudo, essa versão dramatizada da vida do doutorando não foi realidade para mim. Ainda bem. Obviamente tive meus momentos de insegurança, de ansiedade, de pressão, de frustração... Mas não passaram de momentos fortuitos. Tenho certeza, sou uma pessoa privilegiada. As pessoas que me cercaram nesse período minimizaram em muito os percalços. Sou muito grato a elas, e nesta parte da tese quero tentar retribuir um pouco todo suporte delegado a mim.

Primeiramente, sou muito grato à professora Irene. Estou seguro que nem todos têm o privilégio de ter uma orientadora como a minha. Foram muitos anos de trabalho juntos, e sempre tivemos uma relação amistosa. Agradeço pelos ensinamentos. Agradeço muito pela confiança, por ter acreditado naquele guri com dificuldades para fazer *miniprep*. Paciência definitivamente é uma de suas virtudes. Agradeço por acreditar nas minhas ideias e por me dar confiança nas horas que precisei.

Sou também muito grato aos meus pais e ao meu irmão. Sem estrutura familiar, nada de doutorado. Fato. Obrigado, pai, pelo carinho, pelos conselhos, pelos ensinamentos, pelo exemplo de competência e disciplina. És minha referência, meu mentor, meu companheiro. Obrigado, mãe, pelo carinho, pelos conselhos, pelos estímulos, pelas cobranças. Certamente minha escolha profissional foi decorrente do meu fascínio pela natureza, despertado por ti na minha infância. Obrigado, Dudu, pelo carinho e apoio. Meu parceirão!

Sou muito grato a minha namorada Shana. Se eu pudesse dar uma recomendação a quem fosse iniciar um doutorado, diria que arranjasse uma doutoranda para namorar. Tudo fica mais fácil. Pode-se discutir *papers* em uma janta, por exemplo. Brincadeiras à parte, por causa dessa guria meu doutorado foi feliz. Obrigado, Shana, pelo amor, apoio, confiança, amizade. Atravessamos juntos uma graduação, um mestrado e um doutorado, e está cada vez melhor. Acho que vamos ter que casar...

Agradeço muito aos meus avós, primos, tios, sogros e padrinhos. Em especial, gostaria de agradecer meu padrinho/tio João. Obrigado, tio João. Não imagina como foste fundamental na minha infância (não que não continue sendo). Não só as brincadeiras e piadas foram importantes. Percebo agora o quanto estimulaste meu lado científico. Nunca esquecerei os teus presentes criativos, como aquele livro de dinossauros e aquela luneta.

Agradeço a meus amigos de faculdade, de laboratório, de infância, de futebol, de segundo grau. Vocês foram muito importantes para aliviar os momentos de tensão. Valeu!

Sou grato ao pessoal do laboratório, Débora, Dieime, Fran, Maicon, Luciano e Ricardo. Sou particularmente grato à Débora, pois participamos de trabalhos em conjunto, e sua contribuição foi fundamental. Obrigado, Débora, pela amizade. Contigo percebi o que realmente é excelência. Dieime e Maicon, obrigado por confiarem em mim. Aprendi muito com vocês. Ricardo, dou graças a Deus pela tua amizade e ajuda (principalmente quando entrei no lab). Fran e Luciano, obrigado pela amizade e ajuda.

Agradeço ao professor Arnaldo Zaha e à professora Marilene Vainstein pelas sugestões e críticas valorosas. Fico honrado que cientistas desse calibre tenham composto minha comissão de acompanhamento,

Agradeço ao professor Augusto Schrank pelas críticas, pelos ensinamentos e pelas oportunidades concedidas (principalmente em relação ao genoma de *Azospirillum amazonense*). Sem dúvida é uma mente privilegiada, a qual admiro.

Agradeço também ao professor Henrique Ferreira pelos ensinamentos e oportunidades concedidas. Apesar de ele ser colorado, é um exemplo a ser seguido.

Agradeço à Dra. Ana Tereza Ribeiro de Vasconcelos e ao Me. Luiz Gonzaga Paula de Almeida por todo apoio concedido no trabalho de sequenciamento do genoma de *A. amazonense*.

Agradeço também ao professor Arthur Germano Fett, à professora Sídia Jacques, à professora Luciane Passaglia e ao professor Giancarlo Pasquali, que de diferentes formas contribuíram para o meu trabalho.

Sou grato aos funcionários do Cbiot, Sílvia, Luciano e Milton por toda ajuda. Quem dera todos fossem como vocês, prestativos, competentes e simpáticos.

Finalmente, agradeço aos órgãos financiadores FAPERGS, CAPES e CNPq, que possibilitaram a execução deste trabalho.

ÍNDICE

LISTA DE ABREVIATURAS

AIA	ácido indol acético
ACC	aminociclopropano-1-carboxilato
CDS	*coding sequence* (sequência codificadora)
DNA	ácido desoxirribonucleico
EGFP	*enhanced green fluorescent protein* (proteína fluorescente verde reforçado)
EYFP	*enhanced yellow fluorescent protein* (proteína fluorescente amarelo reforçado)
ORF	open reading frame (fase aberta de leitura)
PCR	*polymerase chain reaction* (reação em cadeia da polimerase)
rDNA	DNA ribossômico
RPCVs	Rizobactérias promotoras do crescimento vegetal
RubisCO	Ribulose-1,5-bisfosfato carboxilase oxigenase

RESUMO

A utilização massiva de fertilizantes químicos na agricultura tem efeitos perniciosos ao ambiente. As bactérias do gênero *Azospirillum* são amplamente estudadas, pois são capazes de promover o crescimento vegetal. Essa característica lhes confere potencial para serem utilizadas na agricultura como uma alternativa ecologicamente compatível. Embora a espécie *Azospirillum amazonense* seja menos conhecida, o estudo de sua biologia molecular poderia contribuir para a elucidação dos mecanismos envolvidos na promoção do crescimento vegetal.

Na primeira parte deste trabalho, foram descritas ferramentas genéticas que podem facilitar o estudo da biologia molecular de *A. amazonense*. Métodos de conjugação e eletroporação foram otimizados utilizando vetores com origens de replicação de amplo espectro (pVS1 e pBBR1). Além disso, mutantes para o gene *glnK* foram gerados utilizando o sistema do vetor pK19MOBSACB. Finalmente, um protocolo de análise de promotores baseado na expressão de proteínas fluorescentes foi desenvolvido para permitir estudos de regulação gênica.

Na segunda parte do trabalho, uma análise abrangente das características do *draft* do genoma de *A. amazonense* foi realizada. Essa espécie apresenta um repertório versátil de genes, crucial para seu modo de vida na rizosfera. Genes putativos relacionados com metabolismo de nitrogênio e de carbono, produção de energia, produção de fitormônio, transporte, *quorum sensing*, resistência a antibióticos, síntese de bacteriofitocromo, quimiotaxia e motilidade foram identificados. Os genes da fixação do nitrogênio e da nitrilase poderiam estar diretamente relacionados com a promoção do crescimento vegetal. A identificação de genes da RubisCO sugere que *A. amazonense* seja capaz de fixar carbono, característica do seu metabolismo antes desconhecida. Outro aspecto relevante é que alguns genes de *A. amazonense*, como os da nitrogenase e da RubisCO, são mais próximos filogeneticamente aos genes de membros da ordem Rhizobiales do que dos de espécies do mesmo gênero.

ABSTRACT

The massive use of chemical fertilizers in agriculture has harmful effects to the environment. Bacteria from the *Azospirillum* genus are widely studied, since they are able to promote plant growth. This feature gives them the potential to be used in agriculture as an ecologically compatible alternative. Although the *Azospirillum amazonense* is a lesser-known species, the study of its molecular biology could contribute to a better understanding of the mechanisms implicated in plant growth.

In the first part of this study, genetic tools that can support the study of the molecular biology of *A. amazonense* were described. Conjugation and electrotransformation methods were established utilizing vectors with broad host-replication origins (pVS1 and pBBR1). Furthermore, *glnK*-specific *A. amazonense* mutants were generated utilizing the pK19MOBSACB vector system. Finally, a promoter analysis protocol based on fluorescent protein expression was optimized to aid genetic regulation studies on this bacterium.

In the second part of this study a comprehensive analysis of the genomic features of this species was presented. The species *A. amazonense* presents a versatile repertoire of genes crucial for its plant-associated lifestyle. Genes of *A. amazonense* related to nitrogen/carbon metabolism, energy production, phytohormone production, transport, quorum sensing, antibiotic resistance, chemotaxis/motility and bacteriophytochrome biosynthesis were identified. Noteworthy genes were the nitrogen fixation genes and the nitrilase gene, which could be directly implicated in plant growth promotion, and the carbon fixation genes, which had previously been poorly investigated in this genus. One important finding was that some *A. amazonense* genes, like the nitrogenase genes and RubisCO genes, were closer phylogenetically to genes from Rhizobiales members than to those from species of its own order.

INTRODUÇÃO

Fertilizantes químicos

As plantas necessitam de elementos químicos para sua adequada nutrição. Deficiências na disponibilidade desses elementos podem afetar o crescimento vegetal e, por conseguinte, a produtividade agrícola. Portanto, a utilização de fertilizantes químicos foi um dos alicerces do aumento expressivo na produtividade agrícola mundial, que se iniciou por volta da metade do século passado (HUANG *et al.*, 2002). O aumento populacional progressivo e a consequente demanda por produtos de origem agrícola são os principais responsáveis pelo consumo cada vez maior de fertilizantes. Em relação à quantidade utilizada desses compostos, os fertilizantes nitrogenados compõem a parcela majoritária (cerca de 100 milhões de toneladas ao ano) (HTTP://www.fertilizer.org), já que normalmente a escassez de nitrogênio no solo é o principal fator limitante do crescimento vegetal (RAVEN *et al.*, 2001).

Embora necessário, o uso contínuo e elevado de fertilizantes inorgânicos tem consequências perniciosas para o meio-ambiente. Tendo em vista que esses efeitos são oriundos principalmente da utilização de fertilizantes nitrogenados, os exemplos subsequentes se referem àqueles causados por esses compostos. Segundo Sutton *et al.* (2011), os principais são:

- a diminuição da biodiversidade dos ecossistemas, no qual o fenômeno de eutrofização (multiplicação excessiva de algas) exerce papel principal;

- o risco à saúde humana, pois o consumo de água com altas concentrações de nitrato aumenta a chance de desenvolvimento de câncer de intestino e que a exposição a óxidos de nitrogênio por via aérea causa doenças respiratórias e cardiovasculates;

- e o aumento da concentração de gases do efeito estufa na atmosfera.

Esses problemas ambientais têm consequências econômicas elevadas. Por exemplo, estima-se que o excesso de nitrogênio no ambiente custe à União Européia entre 70 e 320 bilhões de euros ao ano (SUTTON *et al.*, 2011). Portanto, em virtude dos problemas causados pela utilização de fertilizantes químicos, há uma preocupação crescente em se utilizar métodos que promovam uma agricultura sustentável, menos adversa ao ambiente. Nesse sentido, o uso de bactérias promotoras de crescimento vegetal (RPCVs) se destaca

como uma alternativa ecologicamente compatível. Portanto, diferentes laboratórios do mundo vêm estudando esses organismos, tendo como objetivo final seu emprego na agricultura para reduzir o consumo de fertilizantes químicos.

Rizobactérias promotoras de crescimento vegetal (RPCVs)

Rizobactérias promotoras de crescimento vegetal (RPCVs, ou PGPR, do inglês *plant growth promoting rhizobacteria*) são bactérias de vida livre que exercem um efeito benéfico sobre o crescimento de plantas (ANTOUN & PRÉVOST, 2005; KLOEPPER *et al.*, 1989). Visando a sistematização de seu estudo, essa definição originalmente não contemplaria bactérias simbióticas que produzem nódulos radiculares na planta hospedeira, como os rizóbios e espécies do gênero *Frankia* (ANTOUN & PRÉVOST, 2005). Entretanto, como não há homogeneidade na literatura referente a essa definição, ao longo do texto desta tese, sempre que se fizer menção ao acrônimo RPCVs, o conceito original deverá ser considerado.

As RPCVs ocupam a rizosfera, a camada de solo influenciada pela raiz, atraídas por exsudatos que contêm metabólitos vegetais que podem ser utilizados como nutrientes (KLOEPPER *et al.*, 1989). Em contrapartida, as RPCVs exercem efeitos benéficos sobre o crescimento vegetal, que podem ser diretos, através da síntese de substâncias estimuladoras para a planta, ou indiretos, pela proteção da planta contra doenças provenientes do solo, normalmente causadas por fungos (LUGTENBERG & KAMILOVA, 2009; VERMA *et al.*, 2010). Dentre os mecanismos diretos de promoção do crescimento vegetal, destacam-se a fixação biológica do nitrogênio, a síntese de fitormônios, a inibição da síntese de etileno e o aumento da disponibilidade de nutrientes para a planta. (FUENTES-RAMIREZ & CABALLERO-MELLADO, 2005).

A utilização de RPCVs como inoculantes agrícolas possui muitas vantagens em comparação com fertilizantes e pesticidas químicos (BERG, 2009), já que:

- são mais seguros;
- demonstram reduzido dano ambiental e menor risco potencial à saúde humana;
- são eficazes em pequenas quantidades;
- são capazes de se multiplicar, mas são controlados tanto pela planta quanto pelas populações microbianas nativas;
- e decompõem-se mais rapidamente do que os pesticidas químicos convencionais.

Os gêneros mais conhecidos de RPCVs incluem: *Azospirillum*, *Azotobacter*, *Bacillus*, *Burkholderia*, *Enterobacter*, *Klebsiella* e *Pseudomonas*. Alguns gêneros possuem espécies endofíticas, como *Herbaspirillum seropedicae*, *Azoarcus* spp. e *Gluconacetobacter diazotrophicus*, que são capazes de invadir os tecidos vegetais e por isso, estabelecer uma interação mais íntima com a planta (FUENTES-RAMIREZ & CABALLERO-MELLADO, 2005). Como o organismo de interesse neste trabalho é a bactéria *Azospirillum amazonense*, os tópicos seguintes se deterão na caracterização do gênero e da espécie em questão.

O gênero Azospirillum

O gênero *Azospirillum* pertence à classe α-Proteobacteria e compreende bactérias Gram-negativas de vida livre, flageladas e fixadoras de nitrogênio. Esses microrganismos têm distribuição mundial e são normalmente encontrados na rizosfera e nas raízes de diferentes tipos de plantas. Geralmente, os Azospirilla apresentam um metabolismo do nitrogênio e do carbono versátil, que confere adaptabilidade ao competitivo ambiente da rizosfera (STEENHOUDT & VANDERLEYDEN, 2000). Além do mais, essas bactérias são capazes de acumular grânulos de polihidroxibutirato, que em condições adversas podem servir como fonte de carbono e energia (TAL & OKON, 1985).

Muitos estudos demonstraram que as bactérias do gênero *Azospirillum* exercem efeitos benéficos sobre espécies vegetais. No entanto, apesar de mais de três décadas de estudo intensivo terem decorrido, ainda não se sabe o modo exato de ação dessas bactérias sobre as plantas. Entretanto, até o momento, segundo Bashan & de-Bashan (2010), os resultados indicam que:

- a maioria das linhagens de *Azospirillum* pode fixar nitrogênio, mas somente uma fração dele, se alguma, é transferida para planta;

- a maioria das linhagens sintetiza fitormônios, no entanto a transferência para planta é provavelmente limitada;

- o efeito mais perceptível é a modificação na arquitetura radicular da planta, que confere maior absorção de água e nutrientes. Entretanto, apesar desse fenômeno ser atribuído a ação de fitormônios bacterianos, ainda não se sabe se um conjunto desses elementos ou se um único está envolvido no processo;

Em virtude das discussões existentes em relação à fixação do nitrogênio e à produção de fitormônios, mecanismos alternativos de promoção do crescimento vegetal também têm sido estudados. Existem resultados promissores em relação à solubilização de fosfatos, à redução de estresses abióticos e ao controle biológico de patógenos. Entretanto, as evidências ainda são insuficientes para se chegar a uma conclusão definitiva da importância desses mecanismos na estimulação do crescimento de plantas (BASHAN & DE-BASHAN, 2010).

Em contrapartida, a existência de uma miríade de mecanismos possíveis de promoção do crescimento vegetal através da inoculação com *Azospirillum*, levou a Bashan & de-Bashan (2010) a propor a Hipótese dos Mecanismos Múltiplos. Essa conjectura se baseia na premissa de que não haveria um único mecanismo envolvido, mas uma combinação de poucos ou mais mecanismos na estimulação do crescimento vegetal. Os mecanismos poderiam variar em relação às espécies vegetais, às linhagens de *Azospirillum* e as condições ambientais prevalentes durante a interação. Em outras palavras, esse ponto de vista holístico preconiza que seria precipitado atribuir os efeitos da inoculação a um único mecanismo, pois sem a ação de outros mecanismos, talvez eles fossem parciais ou até mesmo não ocorressem.

Seria importante ressaltar que a maioria dos resultados obtidos é decorrente de estudos realizados com a espécie *Azospirillum brasilense* (TARRAND *et al.*, 1978), considerada como organismo-modelo do gênero. Até o momento, outras quinze espécies de *Azospirillum* foram descritas: *A. amazonense* (MAGALHÃES *et al.*, 1983), *Azospirillum canadense* (MEHNAZ *et al.*, 2007), *Azospirillum doebereinerae* (ECKERT *et al.*, 2001), *Azospirillum formosense* (LIN *et al.*, 2011), *Azospirillum halopraeferens* (REINHOLD *et al.*, 1987), *Azospirillum irakense* (KHAMMAS *et al.*, 1989), *Azospirillum largimobile* (DEKHIL *et al.*, 1997), *Azospirillum lipoferum* (TARRAND *et al.*, 1978), *Azospirillum melinis* (PENG *et al.*, 2006), *Azospirillum oryzae* (XIE & YOKOTA, 2005), *Azospirillum palatum* (ZHOU *et al.*, 2009), *Azospirillum picis* (LIN *et al.*, 2009), *Azospirillum rugosum* (YOUNG *et al.*, 2008), *Azospirillum thiophilum* (LAVRINENKO *et al.*, 2010) e *Azospirillum zeae* (MEHNAZ *et al.*, 2007). Portanto, fica claro que a diversidade biológica do gênero vem sendo pouco explorada, bem como o seu potencial biotecnológico.

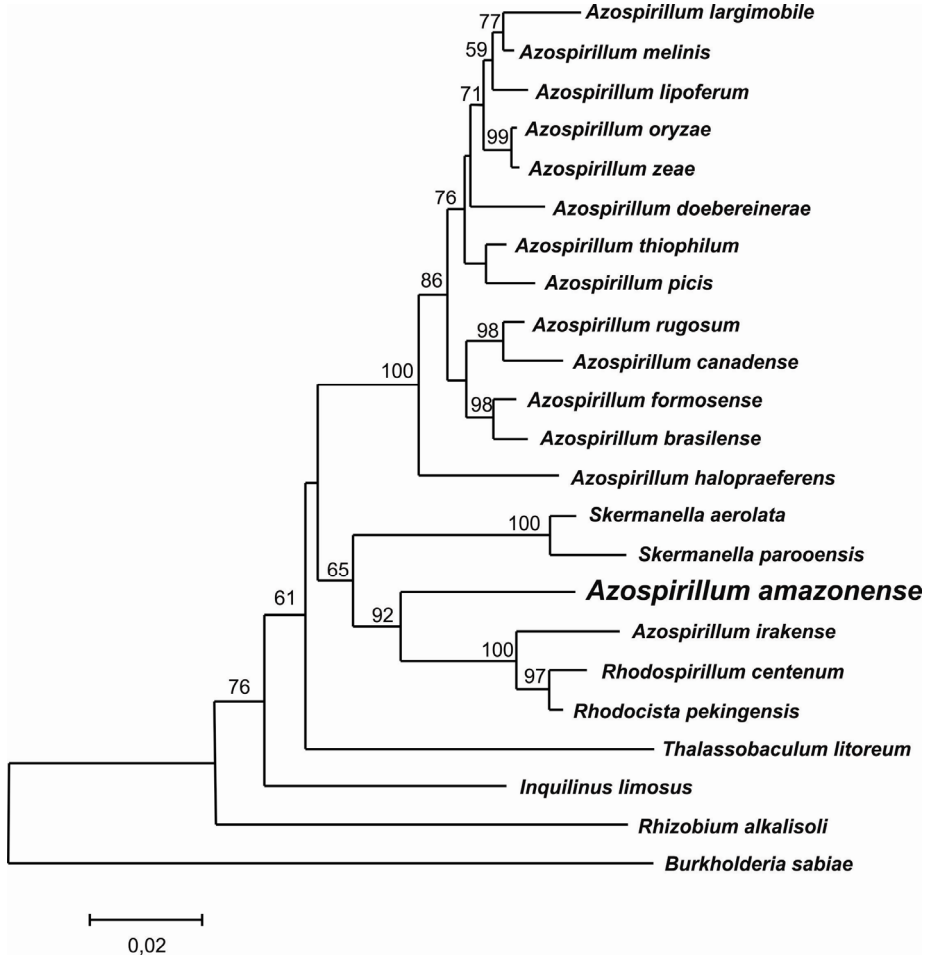


Figura 1. Filogenia de Azospirilla e espécies relacionadas baseada no rDNA 16S.
Esta figura é uma reprodução da análise realizada por Lin *et al.* (2011). O tamanho dos ramos é proporcional à distância evolutiva. Valores de *bootstrap* (1000 réplicas) acima de 50% estão representados adjacentes aos nós.

Além disso, estudos independentes, através da análise filogenética do rDNA 16S (Figura 1), demonstraram que o genêro *Azospirillum* não é um grupo natural (monofilético), mas na verdade um grupo parafilético. Apesar de *A. amazonense* e *A. irakense* pertencerem ao gênero *Azospirillum*, são mais próximos filogeneticamente de espécies dos gêneros *Rhodocista* e *Skermanella* do que de outros Azospirilla propriamente ditos (Figura 1). Por esta razão, estudos concernentes a essas espécies de *Azospirillum* são relevantes, pois podem existir diferenças significativas no modo de ação e na eficácia dessas bactérias sobre o desenvolvimento vegetal.

A espécie Azospirillum amazonense

A espécie *A. amazonense*, alvo do presente trabalho, foi originalmente isolada de gramíneas forrageiras da região Amazônica (MAGALHÃES *et al.*, 1983). Estudos subsequentes demonstraram sua ampla distribuição ecológica, mostrando sua associação também com arroz, sorgo, milho, cana-de-açúcar e outras gramíneas (BALDANI & BALDANI, 2005).

As principais características distintivas dessa espécie em comparação com *A. brasilense*, conforme Magalhães *et al.* (1983) e Fu *et al.* (1989) , são as seguintes:

- o menor tamanho celular;
- a multiplicação preferencial em pH entre 5,8 e 6,8 e a sensibilidade a pH alcalino;
- a utilização de sacarose como fonte de carbono;
- a ausência do sistema DraG/DraT de controle pós-traducional da dinitrogenase redutase;
- a ausência de flagelos laterais.

Embora a bactéria tenha sido descrita em meados dos anos 80, trabalhos sobre sua fisiologia, genética e atividade como RPCV ainda são fortuitos. Recentemente, estudos preliminares têm demonstrado que *A. amazonense* contribui significativamente para o conteúdo de nitrogênio de gramíneas de importância econômica, sendo essa qualidade atribuída a sua capacidade de fixar nitrogênio (REIS JUNIOR *et al.*, 2008; RODRIGUES *et al.*, 2008). Entretanto, estudos adicionais são fundamentais para confirmar esses resultados, para identificar outros mecanismos de ação e para avaliar se a utilização de *A. amazonense* como inoculante seria uma prática ecológica e economicamente viável.

Aspectos gerais da genética molecular de Azospirillum

Diferentes grupos de pesquisa têm feito esforços na investigação de genomas de *Azospirillum*, já que eles poderiam fornecer informações relevantes sobre a biologia desses organismos.

As bactérias desse gênero possuem genomas relativamente complexos, visto que são grandes e são constituídos por múltiplos replicons. O tamanho, o número e a estrutura dos replicons variam conforme as linhagens (MARTIN-DIDONET *et al.*, 2000). Entre as espécies avaliadas, *A. irakense* possui o menor genoma, com 4,8 Mpb distribuídos em

quatro replicons (MARTIN-DIDONET *et al.*, 2000). Em contraposição, o maior genoma, pertencente à bactéria *A. lipoferum* Sp59b, possui 9,7 Mpb distribuídos em oito replicons (MARTIN-DIDONET *et al.*, 2000). Apesar do genoma de *A. amazonense* Y2 ser também constituído por quatro replicons como *A. irakense*, possui menor tamanho, totalizando cerca de 7,3 Mpb (MARTIN-DIDONET *et al.*, 2000). Essa complexidade estrutural dos genomas de *Azospirillum* somada à presença de um número elevado de sequências que codificam transposases são fatores que dificultam projetos genoma envolvendo essas bactérias (VALVERDE *et al.*, 2006). De fato, o único genoma totalmente finalizado do gênero foi o da linhagem *Azospirillum* sp. B510, cuja sequência foi publicada recentemente (KANEKO *et al.*, 2010). Esse genoma é composto por sete replicons de diferentes tamanhos que totalizam cerca de 7,6 Mpb. Além disso, uma característica relevante é a existência de 387 sequências codificantes de transposases (base de dados do KEGG) dispersas nos replicons, observação que ilustra a abundância de sequências repetitivas nos genomas de Azospirilla.

Recentemente, também foi publicado o genoma da espécie *Rhodospirillum centenum* (também conhecida como *Rhodocista centenaria*) (LU *et al.*, 2010), filogeneticamente próxima a *A. amazonense* e *A. irakense* (Figura 1). Em comparação com o genoma de *Azospirillum* sp. B510, o genoma de *R. centenum* é menos complexo, sendo composto por apenas um replicon de 4,35 Mpb e possuindo apenas 38 sequências codificantes de transposases.

O genoma *draft* de *A. brasilense* Sp245 também foi disponibilizado, constituído por 7,5 Mpb divididos em 66 contigs (http://genome.ornl.gov/microbial/abra/19sep08/). Entretanto, por enquanto, uma análise abrangente desses dados ainda não foi publicada.

Embora a disponibilização recente desses genomas tenha sido um progresso fundamental, os dados gerados ainda foram pouco explorados. Até o momento, a maioria dos estudos de caracterização gênica de *Azospirillum* foi realizada com genes isolados de bibliotecas genômicas, se detendo, principalmente, na análise daqueles que em princípio pudessem contribuir para a estimulação do crescimento vegetal, como os envolvidos na fixação do nitrogênio e na síntese de AIA (ácido indolacético, um tipo de auxina). Entretanto, grande parte desses estudos foi realizada com a linhagem *A. brasilense* Sp7, portanto praticamente nada se sabe nesse sentido sobre outras espécies de *Azospirillum*, especialmente *A. amazonense*. Consequentemente, seria relevante sequenciar os genomas

dessas outras bactérias, pois, como relatado anteriormente, a variabilidade interespecífica pode contribuir para descobertas importantes na área de biotecnologia.

Biologia molecular da fixação do nitrogênio

Tendo em vista que o crescimento vegetal parece ser influenciado positivamente por *A. amazonense* através da fixação do nitrogênio, este tópico versará sobre as bases moleculares desse processo.

O nitrogênio é um elemento que constitui uma ampla diversidade de moléculas biológicas, especialmente proteínas e ácidos nucleicos. Embora componha cerca de 80% da atmosfera como nitrogênio molecular (N_2), muitos organismos não são capazes de assimilá-lo nesta forma. Para isso, o nitrogênio molecular deve ser necessariamente convertido a amônia, fonte nitrogenada mais reativa e assimilável. Esse processo fundamental para a disponibilização de nitrogênio para todos os organismos é denominado fixação do nitrogênio e ocorre através de duas vias principais: a industrial, pelo método de Haber-Bosch, e a biológica, pelos organismos diazotróficos (RUBIO & LUDDEN, 2008). Como previamente salientado, a fixação industrial gera consequências adversas ao ambiente, por isso a fixação biológica tem sido estudada.

O processo biológico de fixação é realizado por determinadas espécies procarióticas através de metaloenzimas conservadas denominadas nitrogenases. O tipo mais estudado dessas enzimas é a nitrogenase dependente de molibdênio, que possui dois componentes proteicos principais (RUBIO & LUDDEN, 2008):

- a molibdênio-ferro-proteína ou dinitrogenase, codificada pelos genes *nifD* e *nifK*, que é um complexo heterotetramérico ($\alpha_2\beta_2$) contendo um cofator FeMo no sítio ativo de cada subunidade α (NifD) e um grupamento P em cada interface entre as subunidades α e β;

- e a ferro-proteína ou dinitrogenase redutase, codificada pelo gene *nifH*, que é um complexo homodimérico (γ_2) contendo sítios de ligação a MgATP em cada subunidade e um único grupamento [4Fe4S] na interface entre ambas subunidades.

A dinitrogenase redutase tem a função de reduzir a dinitrogenase, transferindo elétrons do seu grupamento [4Fe4S] em uma reação que depende da hidrólise de MgATP. O grupamento P da dinitrogenase realiza a transferência de elétrons da dinitrogenase

redutase para o cofator FeMo, onde o dinitrogênio é reduzido à amônia (BARNEY *et al.*, 2009). Essa reação pode ser representada como segue:

$N_2 + 8e^- + 16\ ATP + 8\ H^+ \rightarrow 2NH_3 + 16\ ADP + 16\ Pi + H_2$

Tanto em *Azospirillum* spp. quanto em *R. centenum*, os genes envolvidos na fixação do nitrogênio estão organizados sequencialmente em grandes grupamentos gênicos (Figura 2).

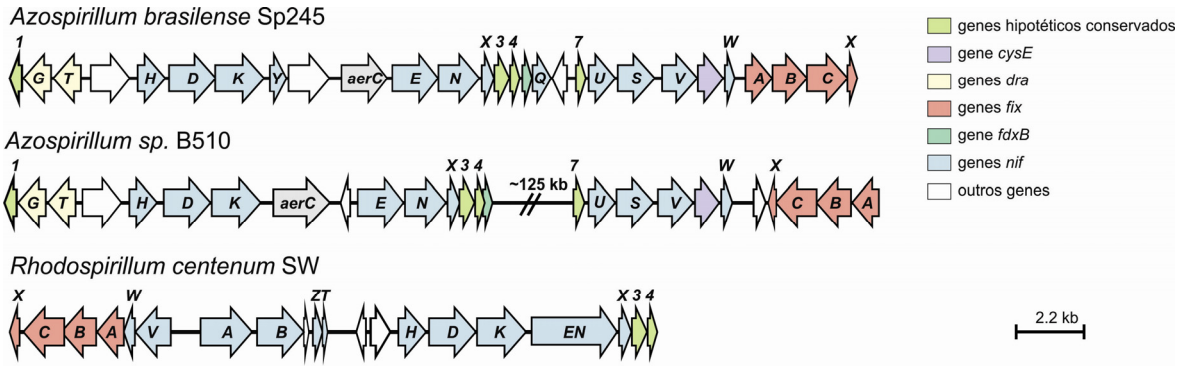


Figura 2. Organização dos genes envolvidos na fixação do nitrogênio em *Azospirillum* spp. e *R. centenum*.
As setas representam os genes e sua respectiva orientação de transcrição. A legenda está representada no canto superior direito da figura.

Seria relevante destacar que o operon *nif* da linhagem *A. brasilense* Sp7 foi o primeiro operon *nif* descrito do gênero (PASSAGLIA *et al.*, 1991; FRAZZON & SCHRANK, 1998; POTRICH *et al.*, 2001; SPEROTTO *et al.*, 2004; SCHRANK *et al.*, 1987), apresentando uma organização muito semelhante ao da linhagem *A. brasilense* Sp245.

Além dos genes *nifHDK*, que codificam os componentes principais da nitrogenase, outros genes presentes nesse bloco são essenciais para a de fixação do nitrogênio. Por exemplo, os genes *nifY*, *nifENX*, *nifUSV* e *nifQ* estão envolvidos na montagem do cofator Fe-Mo, enquanto que os genes *fixABCX*, codificam elementos responsáveis pelo transporte de elétrons para a nitrogenase (RUBIO & LUDDEN, 2008; EDGREN & NORDLUND, 2004).

Tendo em vista que o processo de fixação biológica do nitrogênio é energeticamente dispendioso para célula, existem diferentes mecanismos de controle da expressão e da atividade da nitrogenase. A Figura 3 representa de forma esquematizada a cascata de regulação da fixação do nitrogênio em *A. brasilense*.

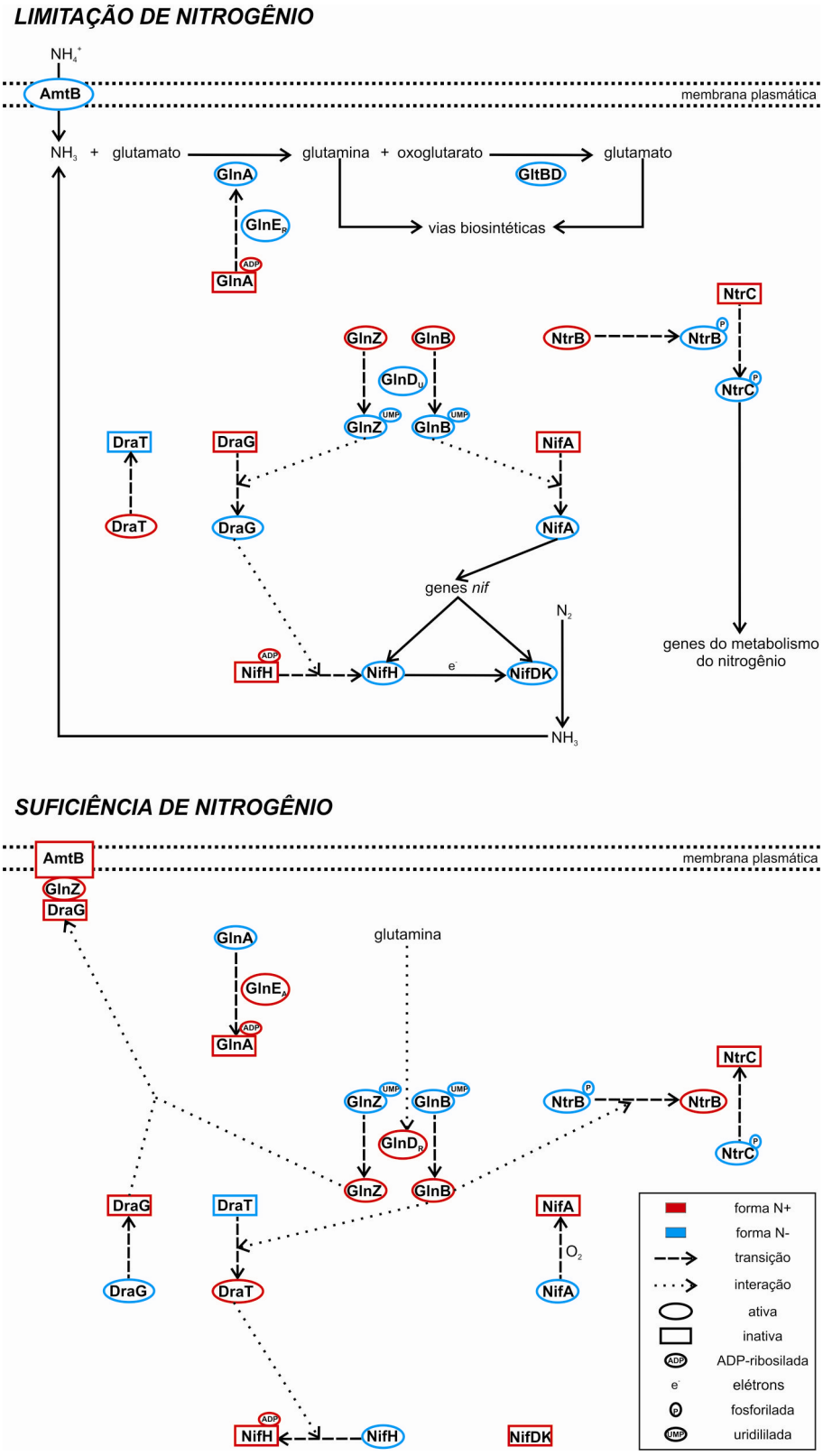


Figura 3. Regulação da fixação do nitrogênio em *A. brasilense.*
A legenda está representada no canto inferior direito da figura. As proteínas PII, GlnZ e GlnB, têm papel central na regulação do metabolismo . A enzima uridililtransferase (GlnD) as modifica pós-traducionalmente

em função da disponibilidade de nitrogênio. Em condições de limitação de nitrogênio, as proteínas PII são uridililadas, e em condições de suficiência de nitrogênio, são desuridililadas. Dependendo de seu estado conformacional, as proteínas PII atuam diretamente e indiretamente na regulação de diferentes alvos proteicos, desde transportadores (AmtB), enzimas (DraG, DraT, NtrB, NifH) e fatores de transcrição (NtrC e NifA).

Essa figura demonstra que, uma vez disponível, a amônia é incorporada ao metabolismo através das enzimas glutamina sintetase (GlnA) e glutamato sintase (GltBD), que produzem os doadores-chave de nitrogênio para as subsequentes reações biosintéticas (WESTBY *et al.*, 1987). Em níveis baixos de glutamina, ou seja, quando o aporte de nitrogênio está reduzido, uma enzima denominada uridililtransferase, codificada pelo gene *glnD*, uridilila as proteínas sinalizadoras PII, que têm papel central na regulação de diversas proteínas do metabolismo nitrogenado (VAN DOMMELEN *et al.*, 2002; ARAÚJO *et al.*, 2008). Em *A. brasilense*, duas proteínas PII parálogas estão presentes, a proteína GlnB e a proteína GlnZ (DE ZAMAROCZY *et al.*, 1996). A proteína GlnB uridililada ativa o fator transcricional NifA, que em baixa tensão de oxigênio, por intermédio do fator sigma N (sigma 54), promove a transcrição de genes relacionados com a fixação do nitrogênio (ARAÚJO *et al.*, 2004; ARSENE *et al.*, 1996; CHEN *et al.*, 2005). Além disso, a proteína GlnB uridililada é incapaz de interagir com a proteína NtrB, que por sua vez fosforila o fator de transcrição NtrC, que ativa a transcrição de genes importantes do metabolismo do nitrogênio, como *glnB* e *glnZ* (DE ZAMAROCZY, 1998; HUERGO *et al.*, 2003). Além disso, GlnB uridililada não interage com DraT, que desta maneira fica inativa, não interferindo no funcionamento da dinitrogenase redutase (HUERGO *et al.*, 2006; HUERGO *et al.*, 2009). Já a proteína GlnZ uridililada ativa a proteína DraG, que remove a ADP-ribose da dinitrogenase redutase, ativando-a (HUERGO *et al.*, 2006; HUERGO *et al.*, 2009). A GlnZ uridililada também libera o transportador AmtB para conduzir amônia para o interior da célula (DE ZAMAROCZY, 1998).

Em contraposição, níveis altos de glutamina indicam suficiência de nitrogênio, então a enzima adenilil-transferase (GlnE) inativa a enzima glutamina sintetase (VAN DOMMELEN *et al.*, 2009). Além disso, a glutamina disponível ativa a atividade removedora de uridilil da proteína GlnD, que então remove os grupos uridilil das proteínas PII (VAN DOMMELEN *et al.*, 2002; ARAÚJO *et al.*, 2008). Consequentemente, a proteína NifA fica inativa e os genes da fixação do nitrogênio não são transcritos. Adicionalmente, a proteína GlnB desuridililada promove a atividade de fosfatase de NtrB, que por sua vez remove o fosfato de NtrC, inibindo a transcrição de genes do metabolismo

do nitrogênio (DE ZAMAROCZY, 1998). Além disso, GlnB não-uridililada ativa DraT, que por sua vez inativa a dinitrogenase redutase por ADP-ribosilação, impedindo que a fixação do nitrogênio ocorra (HUERGO *et al.*, 2006; HUERGO *et al.*, 2009). Já a proteína GlnZ não-uridilada interage com AmtB, impedindo o transporte de amônia, e concomitantemente, sequestra a enzima DraG para a membrana celular, impedindo que a dinitrogenase redutase seja reativada (HUERGO *et al.*, 2006; HUERGO *et al.*, 2007).

No entanto, seria pertinente ressaltar que em *A. amazonense* o sistema DraG-DraT não está presente (FU *et al.*, 1989; HARTMANN *et al.*, 1986). Consequentemente, a sua capacidade de fixação do nitrogênio é pouco afetada pela presença de amônia, característica favorável para a utilização desse microrganismo como inoculante (HARTMANN *et al.*, 1986).

Em *Azospirllum* e outros organismos diazotróficos, o processo de fixação do nitrogênio também é modulado pela disponibilidade de O_2, pois os componentes metálicos da nitrogenase são inativados por oxidação. Recentemente, um estudo demonstrou a presença das proteínas FixLJ e FixK em *A. brasilense*, envolvidas na percepção dos níveis de oxigênio (LI *et al.*, 2010). A proteína FixL é uma histidina quinase que detecta oxigênio através de seu grupo funcional heme. Em condições de baixa tensão de oxigênio, essa proteína se autofosforila e transfere o grupo fosforil ao fator de transcrição FixJ. A proteína FixJ fosforilada ativa a transcrição de outro fator transcricional, denominado FixK, que por sua vez, promove a transcrição do gene da proteína NifA (LI *et al.*, 2010). Como relatado anteriormente, a proteína NifA está envolvida na transcrição dos genes que codificam os componentes da nitrogenase, em condições de limitação de nitrogênio. Nas bactérias *Bradyrhizobium japonicum* e *H. seropedicae*, o fator transcricional NifA também está envolvido na percepção dos níveis de oxigênio (FISCHER *et al.*, 1988; OLIVEIRA *et al.*, 2009). Nesses microrganismos e em *A. brasilense*, essa proteína é composta por três domínios principais (DIXON & KAHN, 2004):

- o domínio C-terminal hélice-volta-hélice, que reconhece os sítios de ligação no DNA;

- o domínio catalítico AAA+, que está envolvido na abertura do complexo de transcrição por intermédio do fator sigma N acoplado à RNA polimerase;

- e finalmente, o domínio N-terminal GAF, que inibe a atividade do domínio catalítico em condições inadequadas para fixação do nitrogênio.

Apesar de pouco compreendida, a sensibilidade de NifA ao oxigênio é atribuída a presença de quatro resíduos de cisteína distribuídos no domínio AAA+ e na região ligadora desse domínio com o domínio hélice-volta-hélice (FISCHER *et al.*, 1988; OLIVEIRA *et al.*, 2009). Uma vez que o fator transcricional NifA de *A. brasilense* também possui esses resíduos de cisteína conservados, acredita-se que eles estejam envolvidos na sua sensibilidade ao O_2 (ARSENE *et al.*, 1996), conforme seus equivalentes em *B. japonicum* e *H. seropedicae*.

A complexidade do sistema de fixação de nitrogênio fica evidente pelo número de proteínas e interações envolvidas nesse processo. Embora os sistemas sejam relativamente conservados entre as proteobactérias diazotróficas, existem variabilidades espécie-específicas em seus componentes. Em relação à *A. amazonense*, as características básicas do sistema de fixação são praticamente desconhecidas. Portanto, estudos que permitam a elucidação das bases moleculares de seu sistema de fixação do nitrogênio são pertinentes.

Ferramentas e técnicas genéticas de manipulação

O sequenciamento de genomas teve impacto significativo na área de bacteriologia, pois tem disponibilizado informações relevantes sobre diferentes aspectos da biologia bacteriana. Em vista disso, uma das preocupações atuais é interpretar a quantidade massiva de dados gerada pelos Projetos Genoma. Portanto, com esse objetivo, é fundamental que sejam desenvolvidas ferramentas e técnicas genéticas espécie-específicas, como vetores de clonagem, metodologias de transferência de DNA, sistemas-repórter e principalmente, técnicas de mutagênese sítio-dirigida. Essas últimas talvez sejam as mais importantes, pois estudos de função gênica dependem da geração de mutantes e da avaliação dos fenótipos resultantes (DAVISON, 2002; SCHWEIZER, 2008).

Embora centenas de bactérias já tenham sido sequenciadas, a utilização dessas metodologias está restrita a um número pequeno de espécies, destacadamente ao organismo-modelo *Escherichia coli* (DAVISON, 2002; SCHWEIZER, 2008). Em relação às bactérias do gênero *Azospirillum*, existe uma carência de ferramentas e técnicas mais sofisticadas. Dois exemplos ilustram bem esse problema. O primeiro é em relação à metodologia padrão de transferência de DNA. Apesar de um protocolo de eletroporação ter sido descrito há mais de 20 anos atrás (VAN DE BROEK *et al.*, 1989), a conjugação é a forma usual de se transferir DNA para *Azospirillum* (HOLGUIN *et al.*, 1999).

O segundo exemplo se refere às metodologias de mutagênese sítio-dirigida. Em geral, a maioria dos mutantes de *A. brasilense* foi gerada pela inserção disruptiva de um cassete de resistência no gene-alvo (exemplos podem ser encontrados em (XIE *et al.*, 2010; LERNER *et al.*, 2009; LERNER *et al.*, 2009), artifício que pode causar efeitos polares nas regiões adjacentes à inserção (LINK *et al.*, 1997). A ausência de metodologias mais elaboradas tem como consequência a falta de eficiência na manipulação genética das espécies de *Azospirillum*, o que afeta diretamente a produtividade no estudo dessas bactérias.

Existem características intrínsecas de cada espécie (ou até mesmo de cada linhagem) que têm de ser consideradas ao se desenvolver as metodologias de manipulação genética. Diferenças na susceptibilidade a antibióticos e na competência de transformação são alguns exemplos (DAVISON, 2002). Portanto, as técnicas e ferramentas genéticas devem ser otimizadas exclusivamente para uma espécie ou linhagem determinada, para que se tenha agilidade na obtenção dos resultados experimentais.

Portanto, com a disponibilização do genoma de *A. amazonense*, o estabelecimento de técnicas e ferramentas específicas se tornou imprescindível para aproveitar toda informação provida pelas sequências, tendo em vista que é uma espécie menos conhecida e que a quantidade de recursos para seu estudo é muito limitada.

Objetivo geral

Tendo em vista o escasso conhecimento da biologia molecular de *A. amazonense*, os objetivos principais deste estudo foram:

- Otimizar ferramentas genéticas para o estudo de *A. amazonense;*
- Identificar aspectos genéticos gerais da biologia de *A. amazonense*.

Objetivos específicos

- Otimizar metodologias de transferência de DNA para *A. amazonense*;
- Otimizar uma metodologia de mutagênese sítio-dirigida para *A. amazonense*;
- Otimizar um sistema-repórter para *A. amazonense*;
- Sequenciar e anotar o genoma de *A. amazonense*;

- Identificar e analisar genes potencialmente envolvidos na fixação de nitrogênio em *A. amazonense*;

- Verificar a história filogenética dos genes *nif* de *A. amazonense*;

- Desenvolver um modelo do metabolismo do nitrogênio de *A. amazonense*;

- Identificar genes no genoma de *A. amazonense* potencialmente fundamentais a seu modo de vida adaptado à interação com plantas.

CAPÍTULO 1

Artigo publicado na revista BMC Microbiology (Fator de impacto: 2,96)

Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*

Fernando H Sant'Anna, Dieime S Andrade, Débora B Trentini, Shana S Weber & Irene S Schrank.

BMC Microbiol. 2011; 11: 107

O material suplementar desse trabalho foi removido da versão final do artigo, por isso não há menção sobre ele no manuscrito. Entretanto, ele foi agregado ao final desse capítulo a fim de facilitar o entendimento do trabalho.

BMC Microbiology

METHODOLOGY ARTICLE Open Access

Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*

Fernando H Sant'Anna[1], Dieime S Andrade[1], Débora B Trentini[1], Shana S Weber[1] and Irene S Schrank[2*]

Abstract

Background: *Azospirillum amazonense* has potential to be used as agricultural inoculant since it promotes plant growth without causing pollution, unlike industrial fertilizers. Owing to this fact, the study of this species has gained interest. However, a detailed understanding of its genetics and physiology is limited by the absence of appropriate genetic tools for the study of this species.

Results: Conjugation and electrotransformation methods were established utilizing vectors with broad host-replication origins (pVS1 and pBBR1). Two genes of interest - *glnK* and *glnB*, encoding PII regulatory proteins - were isolated. Furthermore, *glnK*-specific *A. amazonense* mutants were generated utilizing the pK19MOBSACB vector system. Finally, a promoter analysis protocol based on fluorescent protein expression was optimized to aid genetic regulation studies on this bacterium.

Conclusion: In this work, genetic tools that can support the study of *A. amazonense* were described. These methods could provide a better understanding of the genetic mechanisms of this species that underlie its plant growth promotion.

Background

Many of the negative ecological impacts of agriculture originate from the high input of fertilizers. The increase of crop production in the future raises concerns about how to establish sustainable agriculture; that is, agricultural practices that are less adverse to the surrounding environment [1,2]. The use of microorganisms capable of increasing harvests is an ecologically compatible strategy as it could reduce the utilization of industrial fertilizers and, therefore, their pollutant outcomes [1,3].

Azospirillum is a well-known genus that includes bacterial species that can promote plant growth. This remarkable characteristic is attributed to a combination of mechanisms, including the biosynthesis of phytohormones and the fixation of nitrogen, the most intensively studied abilities of these bacteria [4]. The species *Azospirillum amazonense* was isolated from forage grasses and plants belonging to the Palmaceae family in Brazil by Magalhães et al. (1983) [5], and subsequent works demonstrated its association with rice, sorghum, maize, sugarcane, and *Brachiaria*, mainly in tropical countries [6]. When compared with *Azospirillum brasilense*, the most frequently studied species of the genus, *A. amazonense* has prominent characteristics such as its ability to fix nitrogen when in the presence of nitrogen [7] and its better adaptations to acidic soil, the predominant soil type in Brazil [5,8]. Moreover, Rodrigues et al. (2008) [8] reported that the plant growth promotion effect of *A. amazonense* on rice plants grown under greenhouse conditions is mainly due to its biological nitrogen fixation contribution, in contrast to the hormonal effect observed in the other *Azospirillum* species studied.

Despite the potential use of *A. amazonense* as an agricultural inoculant, there is scarce knowledge of its genetics and, consequently, its physiology. Currently, the genome of *A. amazonense* is being analyzed by our group and its completion will be forthcoming; therefore, the development of specific genetic tools is crucial for taking full benefit of the data. that will be generated.

* Correspondence: irene@cbiot.ufrgs.br
[2]Departamento de Biologia Molecular e Biotecnologia - Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (Av. Bento Gonçalves, 9500, Campus do Vale), RS, Brazil
Full list of author information is available at the end of the article

Hence, in this work we describe methods for the genetic manipulation of *A. amazonense*: DNA transfer methodologies (conjugation and electroporation), reporter vectors, and site-directed mutagenesis. In order to demonstrate the applicability of the optimized techniques, we show the results obtained in the study of the PII signaling proteins of *A. amazonense*, starting from their gene isolation.

Results and Discussion

Isolation of *glnB* and *glnK* genes from *A. amazonense*

The PII proteins are pivotal regulators of the nitrogen metabolism, controlling the activities of transporters, enzymes and transcriptional factors implicated in this process [9,10]. These proteins are highly conserved and are widely distributed throughout prokaryotes [11]. In Proteobacteria in particular, there are two main types of PII proteins, GlnB and GlnK. In this work, two PII protein encoding genes from *A. amazonense* were isolated. Southern blot analysis utilizing a PCR-generated *glnB* fragment as the probe revealed two distinct signals in the genomic DNA of *A. amazonense* digested with SalI: the strongest at the ~2 kb DNA fragments and the weakest at the ~3 kb DNA fragments (data not shown). Based on these results, a genomic library enriched with 2-3 kb SalI fragments was constructed. The library was partially sequenced and a PII protein homolog was identified. The deduced amino acid sequence of this gene was found to be highly similar to that of the GlnZ proteins (GlnK-like homologs) from *A. brasilense* and *Azospirillum* sp. B510 (75% identity and 86% similarity), and *Rhodospirillum. centenum* (73% identity and 86% similarity). Arcondéguy et al. (2001) [12] suggested that the *glnZ* genes should be termed *glnK*, since their deduced proteins are highly similar to the GlnK proteins. Furthermore, there is a functional correspondence between these proteins, as both regulate the uptake of ammonium through the AmtB transporters [13-15]. Therefore, we adopted the *glnK* designation for this *A. amazonense* homolog, mainly because this nomenclature could facilitate comparisons between other bacterial systems.

The *glnK* gene from *A. amazonense* is flanked by the *aat* gene in the downstream region, which codes a putative aspartate aminotransferase and the *ubiH* gene in the upstream region, which codes an enzyme implicated in ubiquinone biosynthesis (Figure 1). This genetic organization resembles that found in other species from the Rhodospirillales order, namely *A. brasilense*, *Azospirillum* sp. B510 and *R. centenum*.

Since the *glnB* gene was not found in the genomic library, the Inverse PCR methodology was carried out to isolate this gene. A ~2 kb amplicon that contained the *glnB* gene was obtained (data not shown). It was found that the protein of this gene displays 92% identity and 98% similarity to the GlnB proteins from *Azospirillum* sp. B510 and *A. brasilense*, and 96% identity and 98% similarity to the GlnB protein of *R. centenum*. The *glnB* gene is located upstream of the *glnA* gene (glutamine synthetase), the same genetic context observed in these bacteria (Figure 1).

In *A. brasilense*, *glnB* has a key role in nitrogen fixation because its protein product regulates the activity of NifA, the transcriptional factor of nitrogen fixation [16,17].

Furthermore, both of the GlnZ (GlnK-like homolog) and GlnB proteins are also implicated in the DraT/DraG system, which regulates dinitrogenase reductase activity by covalent modifications [15]. However, Fu et al. [18] verified that *A. amazonense* does not have the DraT/DraG system. Hence, in the near future, the interaction targets of the PII protein in *A. amazonense* should be determined to better understand their roles in the nitrogen metabolism of this microorganism.

Antibiotic minimum inhibitory concentration

Most DNA manipulation is dependent on the use of vectors containing resistance markers to antibiotics [19,20]. In a previous work using antibiotic susceptibility test discs, Magalhães et al. (1983) [5] showed that *A.*

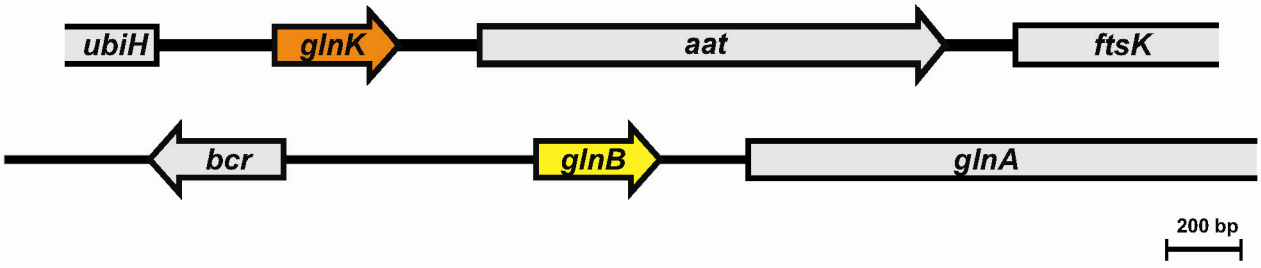


Figure 1 Physical maps of the *glnK* and *glnB* regions of *A. Amazonense*. Genes are represented by the large arrows; *glnA*, *ubiH* and *ftsK* were not completely sequenced.

amazonense is sensitive to kanamycin and gentamicin, tolerant to tetracycline, and resistant to penicillin. In this work, we determined the minimum inhibitory concentrations of *A. amazonense* to antibiotics that are normally used to provide a selective pressure for vectors.

The susceptibility of *A. amazonense* to kanamycin and gentamicin was confirmed, since no growth was observed in concentrations of these antibiotics of 0.25 μg/mL; therefore, vectors that contain selection markers for these compounds are appropriate for use.

High concentrations of ampicillin (128 μg/mL) were required for complete growth inhibition, showing that *A. amazonense* is also resistant to this beta-lactam antibiotic.

It is worth noting that the growth of *A. amazonense* was absent in a relatively high concentration of tetracycline (32 μg/mL), indicating that this species is, in fact, resistant to this antibiotic, instead of tolerant, as pointed out by Magalhães et al. [5]. These findings about the latter two antibiotics are relevant because they could be used in counter-selection procedures in conjugation experiments, as there is a variety of *E. coli* strains that are susceptible to them.

Conjugation

Conjugation mediated by *E. coli* is the standard DNA transfer technique of the *Azospirillum* genus [21]. Therefore, in this work the conjugation ability of *A. amazonense* was evaluated.

Unlike *A. brasilense*, *A. amazonense* cannot grow in LB medium. Furthermore, *E. coli* cannot grow in M79 medium; therefore, the first concern was to establish a medium that provided appropriate growth conditions for the donor and recipient strains. Hence, different medium compositions, containing distinct ratios of M79 and LB media (varying from 1:1 to 9:1), were prepared. The medium mixture of M79:LB at a proportion of 8:2 was the most suitable for culturing both bacteria and it was designated as MLB medium.

Another requisite for the conjugation procedure is to select vectors that contain proper selection markers that are mobilizable and able to replicate inside the receptor cell [19,20]. Therefore, the pHRGFPGUS (pBBR1 replication origin) and the pPZPLACEYFP (pVS1 replication origin) plasmids were tested by tri-parental conjugation. These plasmids are mobilizable broad-host vectors harboring kanamycin resistance markers and fluorescent protein coding genes, which could promptly report achievement of the DNA transfer. The transconjugants exhibited kanamycin resistance and fluorescence. The conjugation frequencies were 3.8×10^{-8} per recipient cell for the pHRGFPGUS vector and 3.8×10^{-7} for the pPZPLACEYFP vector.

Different ratios of recipient to donor and helper strains (1:1:1, 5:1:1, 10:1:1 and 20:1:1) were also tested. The best efficiencies were obtained with the ratios 10:1:1 and 5:1:1; however, no obvious differences between these latter ratios were observed (data not shown).

In conclusion, conjugation is an appropriate method for DNA transfer to *A. amazonense*. Although only tri-parental mating was tested in this work, it is important to mention that bi-parental conjugation could be an alternative test, due to the possibility of increasing the conjugation efficiencies.

Electrotransformation

Since suitable vectors for *A. amazonense* were defined and since conjugation is a time-consuming procedure, the transformation of *A. amazonense* via electroporation was tested.

The eletrocompetence of the cells is greatly influenced by the growth phase [22]. Therefore, *A. amazonense* cells were harvested at different growth phases to evaluate their effect on electroporation efficiency. Cells from the late-log phase (OD_{600} 1) and the stationary phase (OD_{600} 2) were not electrocompetent. Electroporation utilizing cells from the early-log growth phase (OD_{600} 0.12) generated a significant number of transformants. Therefore, all subsequent tests were performed utilizing cells cultivated at this growth phase.

In the electrocompetent cell preparation, the cells were harvested and washed continuously until the solution had a low-ionic strength. The $MgCl_2$ HEPES-sucrose buffer was found to be the most suitable solution for the preparation of *A. amazonense* electrocompetent cells. Although 10% glycerol solution is commonly used for electrocompetent cell preparation in a diverse number of species (including *A. brasilense*), it was not appropriate for *A. amazonense*, as no transformants were obtained when this solution was used.

Different electroporation parameters were tested. The increase in electrical field strength had a positive effect on electroporation efficiency (Figure 2A). The highest electrical field strength tested was 12.5 kV/cm, and this condition was found to be the most efficient, generating about 8000 transformants/μg of pHRGFPGUS (Figure 2A). The effect of pulse length on electroporation efficiency was also investigated (Figure 2B). A pulse length of 4.3 ms (electroporation apparatus set at 200Ω) was the most efficient. The pulse lengths of 7.3 ms (400 Ω) and 10.5 ms (600 Ω) had a dramatic negative effect on transformation efficiency, where only few transformants were obtained (Figure 2B). These conditions are in agreement with the general parameters of bacterial electroporation [22-24].

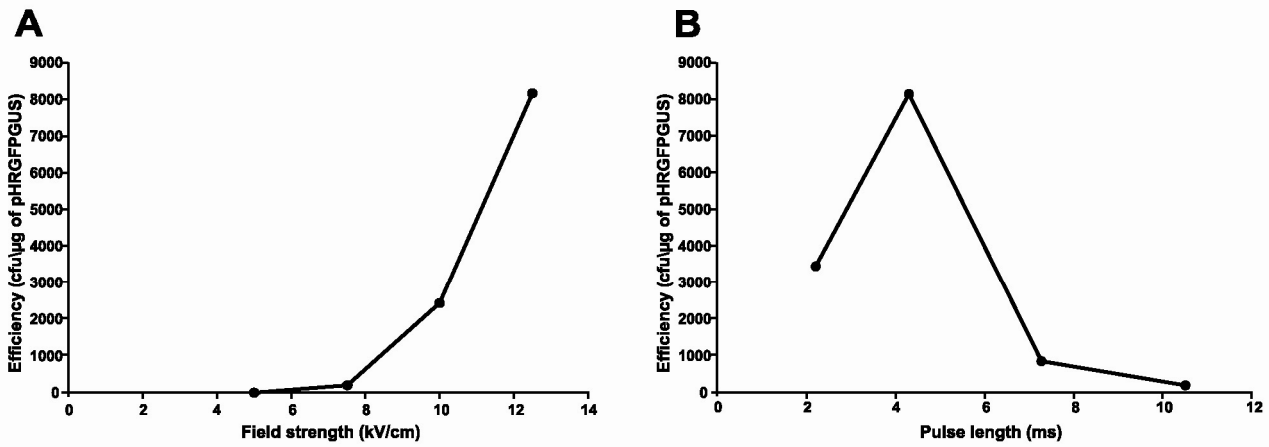


Figure 2 Electrical parameters tested for the *A. amazonense* electroporation. A - Effect of electrical field strength on the transformation efficiency of *A. amazonense*. Competent cells were electroporated at the electric field strengths indicated with the pHRGFPGUS vector, with the GenePulser apparatus set at 200 Ω and 25 μF. B - Effect of the pulse length on the transformation efficiency of *A. amazonense*. Competent cells were electroporated with different pulse lengths, using 50 ng of the pHRGFPGUS vector and with the GenePulser apparatus set at 12.5 kV/cm and 25 μF. The pulse lengths 2.2 ms, 4.3 ms, 7.3 ms and 10.5 ms are obtained setting the GenePulser apparatus at 100 Ω, 200 Ω, 400 Ω and 600 Ω, respectively.

In conclusion, the transfer of DNA to *A. amazonense* by means of electroporation was demonstrated. Although the efficiency of electrotransformation was far from desirable, this result is supported by previous works showing that bacteria closely related to *A. amazonense*, such as *A. brasilense* [25], *R. rubrum* [26] and *Magnetospirillum gryphiswaldense* [27], are recalcitrant to electrotransformation. Nonetheless, this technique is an easy and a rapid method of DNA transfer to the cells of *A. amazonense*.

Site-directed mutagenesis

Site-directed mutagenesis is a fundamental tool for correlating cellular functions with specific regions of the DNA. Therefore, once DNA transfer techniques were established for *A. amazonense*, the next step was to determine a site-directed mutagenesis protocol for this species.

Most of the *A. brasilense* mutants have been generated by the disruptive insertion of an antibiotic resistance cassette into the target gene [14,28-30]. This approach is not recommended when the target gene composes an operon, since the resistance cassette could introduce a polar effect on the expression of the surrounding genes and, consequently, make it difficult to assign a mutant phenotype to the disrupted gene [31].

Therefore, in this work, a site-directed mutagenesis methodology that generates in-frame mutants without the disruptive insertion of a resistance cassette was evaluated. The *glnK* gene was selected for this methodology because subsequent studies of our laboratory will aim to determine the role of the PII proteins in *A. amazonense* metabolism.

The mutagenesis methodology is depicted in Figure 3A. Firstly, an amplicon containing an in-frame deletion of the *glnK* gene was generated through Crossover PCR, and it was subsequently cloned in the suicide replacement vector pK19MOBSACB, generating the pKΔK plasmid. This vector contains a kanamycin resistance gene (positive selection marker) that allows the selection of bacteria that would have integrated the plasmid into the chromosome. This vector was delivered to *A. amazonense* by means of conjugation (the carbon source utilized was maltose instead of sucrose) and one colony resistant to kanamycin was obtained, suggesting that the integration of the plasmid was successfully accomplished. The *sacB* gene (negative marker selection) of the vector is lethal in the presence of sucrose; therefore, the merodiploid strain (containing both wild-type and mutant alleles) was unable to grow in M79 (containing 10 g/L of sucrose). Subsequently, expecting that a recombination event could replace the wild-type allele, the merodiploid strain was cultured for many generations in M79 containing maltose instead of sucrose. Finally, this culture was plated in M79 containing sucrose to eliminate the bacteria that did not accomplish the second recombination event. Seven sucrose-resistant/kanamycin-sensitive colonies were chosen for PCR evaluation of the substitution of the mutant allele for the wild-type gene. Four colonies presented a band of 121 bp, indicating that the wild-type *glnK* was successfully substituted, whereas three colonies presented the 361 bp band, corresponding to the wild-type allele (Figure 3B). Furthermore, an additional PCR with primers flanking the recombination sites was performed, and it also demonstrated a reduction of the amplicon sizes originated from the *glnK* mutants in relation to the

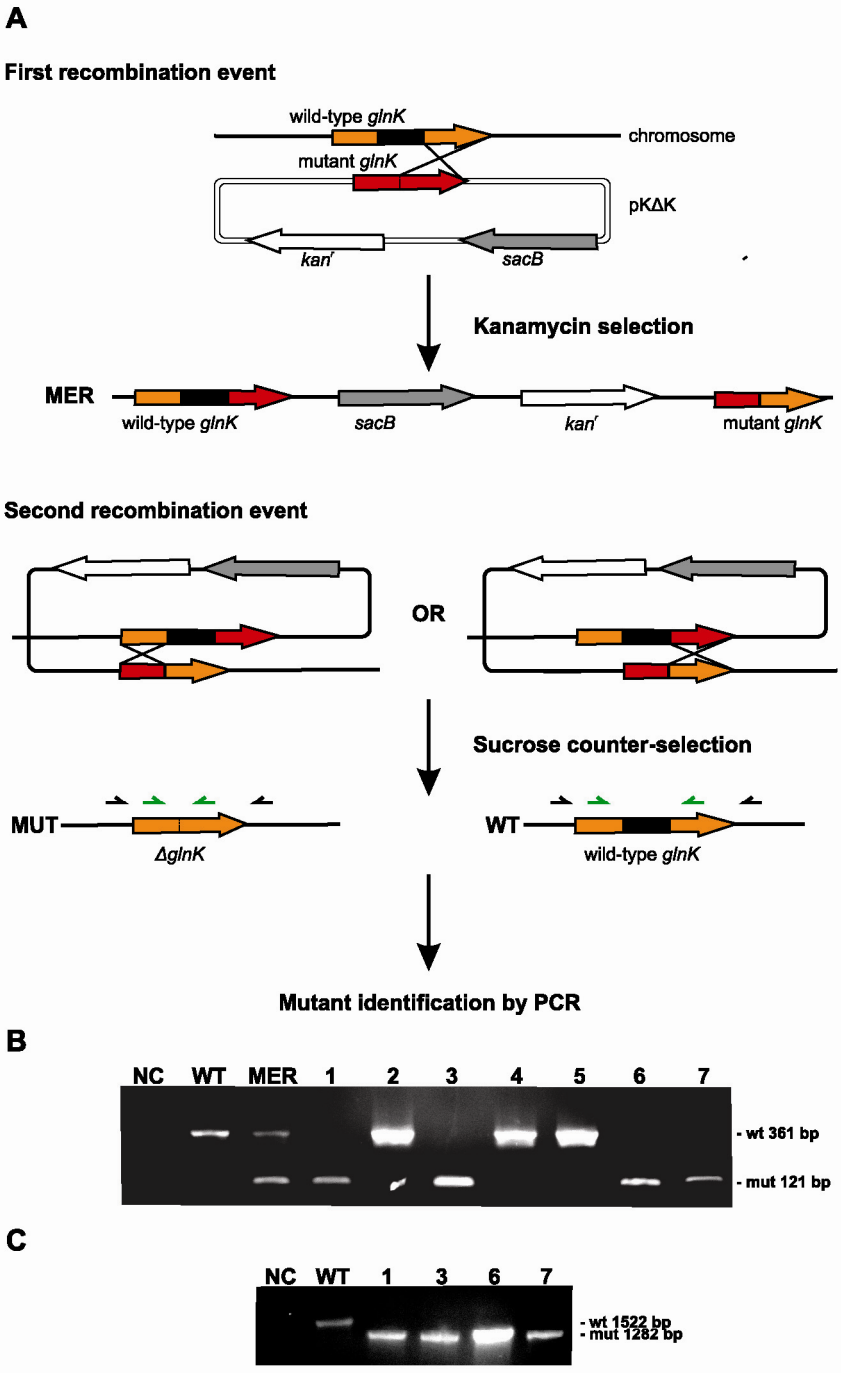


Figure 3 *glnK* gene mutagenesis. A - Schematic diagram depicting the mutagenesis procedure (modified from Clerico et al., 2007 [42]). The vector pKΔK (pK19MOBSACB derivative) harbors the flanking regions of the *glnK* gene (red). This suicide plasmid was delivered by conjugation to *A. amazonense* and integrated in the target site (orange) by homologous recombination, generating a merodiploid strain (containing both, wild-type and mutant alleles) that was selected by kanamycin since there is a resistance marker (white) present in the vector. The black box represents the region deleted. Subsequently, the merodiploid strain was cultivated and the cells that underwent a second recombination event were selected by sucrose, since the *sacB* marker present in the vector is lethal in the presence of this substance. The kanamycin-sensitive/sucrose resistant colonies were evaluated by PCR. B - Identification of the mutant strains by PCR using primers that flank the deletion site. The primers glnK_NdeI_up and glnK_BamHI_do utilized in this procedure are represented by the small green arrows in Figure 3A. NC - negative control, WT - wild type, MER - merodiploid, numbers - strains tested. C - Verification of the mutant strains by PCR using primers that flank the recombination sites. The primers conf_glnK_up and conf_glnK_do are represented by the small black arrows in Figure 3A. NC - negative control, WT - wild type, numbers - strains tested.

wild type strain (Figure 3C). This latter result demonstrates that recombination occurred in the target site.

Altogether, these results show that an in-frame *glnK* gene mutant strain of *A. amazonense* was successfully generated by this mutagenesis system.

Reporter gene system

The study of promoters is fundamental to elucidation of the genetic regulatory mechanisms of bacterial species. Up until now, there has been neither a report of heterologous gene expression in *A. amazonense*, nor a reporter system designed for this species. In this work, a reporter system based on expression of the Enhanced Yellow Fluorescent Protein (EYFP) was developed to analyze the regulatory regions of *A. amazonense* genes *in vivo*.

In silico analysis using a *Sinorhizobium meliloti* sigma 70 promoter weight matrix revealed that the genes *aat*, *glnK*, and *glnB* of *A. amazonense* have putative promoter sequences in their upstream regions (Figure 4). In *E. coli*, sigma 70 is considered to be the vegetative sigma factor, as it is responsible for the expression of the majority of genes [32,33]. Therefore, one could expect that these putative *A. amazonense* sigma 70 promoters could act under standard laboratory growth conditions (aerobic environment, 35°C and M79 medium). Consequently, different vectors were constructed to determine the activity of the upstream regulatory sequences of *A. amazonense* genes in the expression of EYFP.

The *lac* promoter was utilized as a positive control since there is a report showing that this promoter has high activity in *A. brasilense* [34]. Two different vectors were constructed with the *lac* promoter, one derived from pPZPLACEYFP (pVS1 replicon) and the other derived from pHRGFPGUS (pBBR1 replicon). The upstream regions of the genes *glnB*, *glnK*, and *aat* were cloned into the pHRGFPGUS derivative.

Promoter	Sequence	Score
S. meliloti	CTTGACWNNNNNNNNNNNNNNNNCTATATC	
lac	CTTTACACTTTATGCTTCCGGCTCGTATGTT	8.77
glnK	CTTGTGAATACGG-CACGCTTCTTGAAGAAC	6.24
aat	CTTGTGGGGGCGGGGCCCGGCGGCTAATATC	5.83
glnK	GTTAACCAGGCCT-GGATTGCCTATAATTTA	5.74
aat	CTTGTGGGGGCGG-GGCCCGGCGGCTAATAT	5.55
aat	ATCGAGTTTCTAAGCATTTAGGGGCAAATCC	5.39
glnB	CCTGCCCCGTCAAGCATTCTGATAAAATAAT	5.35
aat	ATCGAGTTTCTAA-GCATTTAGGGGCAAATC	5.19
lac	GTTGTGTGGAATT-GTGAGCGGATAACAATT	5.14
glnK	CTTGTCCAGTGCCACGCTCATGCCTGTTAAC	4.89
W/P	CTTGTCCAGATAGCCCAGTAGCTGACATTCA	4.59

Figure 4 *In silico* sigma 70 promoter analysis. The upstream sequences of the genes were analyzed by Patser software using an *S. meliloti* sigma 70 factor weight matrix [33]. *aat* - upstream region of the *aat* gene; *glnB* - upstream region of the *glnB* gene; *glnK* - upstream region of the *glnK* gene; *lac* - *lac* promoter; W/P - negative control, 500 bp upstream of the *eyfp* gene of the plasmid pHREYFP. The *S. meliloti* promoter consensus is the first sequence. Nucleotides that match the *S. meliloti* consensus are in red, and those that match the most conserved residues of the *S. meliloti* promoter consensus (relative frequencies above 0.8) are in bold. Gaps were inserted to preserve the alignment at the regions of the promoters.

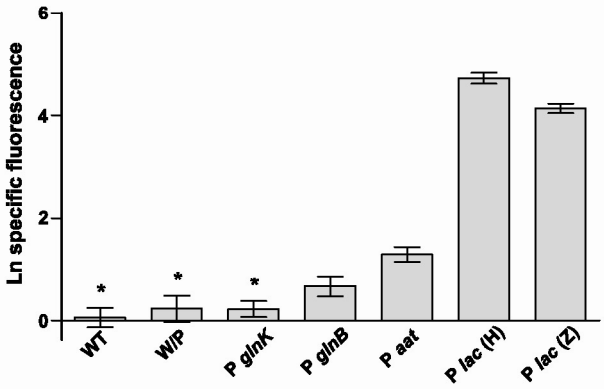


Figure 5 Analysis of EYFP expression controlled by different *A. amazonense* promoters. WT- *A. amazonense* without plasmid; W/P - negative control, *A. amazonense* harboring the pHREYFP vector (without promoter); P *glnK* - *A. amazonense* harboring the pHRPKEYFP vector (promoter of *glnK* gene); P *glnB* - *A. amazonense* harboring the pHRPBEYFP vector (promoter of *glnB* gene); P *aat* - *A. amazonense* harboring the pHRAATEYFP vector (promoter of *aat* gene); P *lac* (Z) - *A. amazonense* harboring the pPZPLACEYFP vector (*lac* promoter); P *lac* (H) - *A. amazonense* harboring the pHRLACEYFP vector (*lac* promoter). The error bars represent the confidence interval of 95%, calculated from seven independent experiments (excepting the P *lac* (H), where four experiments were performed). Asterisks indicate activities that do not differ statistically in the Tukey HSD test ($P < 0.01$).

The *lac* promoter had the best score in the *in silico* analysis from among the promoters detected, and, as expected, the highest fluorescence levels were observed in the *lac* constructions (Figure 5). The difference in the fluorescence levels between the pHRLACEYFP and pPZPLACEYFP transformants could be a product of the difference in the copy number between these vectors.

Although the *in silico* analysis revealed that the *glnK* promoter had a higher score than the *aat* and *glnB* promoters, its *in vivo* activity under the conditions tested did not differ significantly from the negative controls (without promoter and without plasmid) (Figure 5). One of the possible reasons for this is that this gene was repressed under these conditions. The reporter gene analysis also demonstrated that the *aat* and *glnB* promoters were active under the conditions tested, although the *aat* promoter showed a higher activity than the *glnB* promoter.

These observations show that a reporter system based on EYFP can be used for *in vivo* promoter analyses in *A. amazonense*.

Conclusions

Genetic manipulation is fundamental for taking full advantage of the information generated by DNA sequences [20]. Thus, in the present work, we described a series of tools that could assist genetic studies of the diazotrophic bacteria *A. amazonense*, a microorganism presenting potential for use as an agricultural inoculant.

Methods

Bacterial strains, plasmids, and growth conditions

The strains and plasmids utilized in this work are listed in Table 1.

Azospirillum amazonense was cultured in M79 medium (10 g/L of sucrose as the carbon source, 0.1 g/L of K_2HPO_4, 0.4 g/L of KH_2PO_4, 0.2 g/L of $MgCl_2.7H_2O$, 0.1 g/L of NaCl, 0.4 g/L of yeast extract, pH 6.5) [35] at 35°C (unless stated otherwise). The M79 agar plates contained 2.5 mg/L of Bromothymol Blue. *Escherichia coli* XL1-Blue was cultured in LB medium at 37°C [36].

DNA techniques

Standard DNA techniques such as PCR, plasmid extraction, DNA restriction and modification, gel electrophoresis, and *E. coli* transformation were carried out as

Table 1 The strains and plasmids utilized in this work

Bacterial strains	Genotype	Reference	
Escherichia coli XL1-Blue	*recA1, endA1, gyrA96, thi-1, hsdR17(rK-, mK+), supE44, relA1, lac, [F', proAB, laclqZΔM15::Tn10(tet^r)]*	Stratagene	
Azospirillum amazonense Y2	wild type	EMBRAPA-RJ	
Azospirillum amazonense delK	Y2 derivative, *ΔglnK*	This work	
Plasmids	**Relevant characteristics**	**Reference**	**Purpose**
pUC18	cloning vector, amp^r, *lacZα*	Clontech	cloning procedures
pAAGLNK	pUC18 derivative containing the SalI genomic fragment of the *glnK* region	This work	*glnK* isolation
pGLNBA	pUC18 derivative containing the amplification product generated from the revsf_glmBint and gln_AA_do primers	This work	*glnB-glnA* intergenic region and partial *glnA* isolation
pRK2013	ColE1 ori, kan^r, *mob*, *tra*	[45]	helper plasmid for conjugation experiments
pHRGFPGUS	pBBR1 ori, kan^r, *mob*, *gfpmut3*	[46]	DNA transfer evaluation, reporter vectors construction
pPZP201BK	pVS1 ori, ColE1 ori, kan^r, *mob*	[47]	construction of pPZPLACEYFP
pEYFP	pUC ori, *eyfp*, amp^r	Clontech	contruction of reporter plasmids
pPZPLACEYFP	PvuII-EcoRI fragment from pEYFP (containing the *lac* promoter-*eyfp* gene fusion) cloned into the EcoRV-EcoRI sites of pPZP201BK	This work	DNA transfer evaluation, positive control in the reporter assay
pK19MOBSACB	Integration vector, kan^r, ColE1 replication origin, *mob*, *sacB*, *lacZα*	[48]	*glnK* mutagenesis
pKΔK	pK19MOBSACB derivative containing the flanking regions of the *glnK* gene joined by Crossover PCR	This work	*glnK* mutagenesis
pAATEYFP	pEYFP derivative containing the BglII-NcoI fragment of pAAGLNK cloned into its NcoI-BamHI sites	This work	construction of pHRAATEYFP
pPBEYFP	pEYFP derivative containing the intergenic region between the bcr protein superfamily gene and the *glnB* gene cloned into its NcoI-BamHI sites	This work	construction of pHRPBEYFP
pPKEYFP	pEYFP derivative containing the intergenic region between the *ubiH* and the *glnK* genes cloned into its NcoI-BamHI sites	This work	construction of pHRPKEYFP
pHRAATEYFP	HindIII-EcoRI fragment from pAATEYFP (containing the *aat* promoter-*eyfp* gene fusion) joined with the 5.8 kb HindIII-EcoRI fragment of pHRGFPGUS	This work	promoter evaluation
pHRLACEYFP	PvuII-EcoRI fragment from pEYFP (containing the *lac* promoter-*eyfp* gene fusion) joined with the 7.4 kb EcoRV-EcoRI fragment of pHRGFPGUS	This work	promoter evaluation
pHRPBEYFP	HindIII-EcoRI fragment from pPBEYFP (containing the glnB promoter-eyfp gene fusion) joined with the 5.8 kb HindIII-EcoRI fragment of pHRGFPGUS	This work	promoter evaluation
pHRPKEYFP	HindIII-EcoRI fragment from pPKEYFP (containing the *glnK* promoter-*eyfp* gene fusion) joined with the 5.8 kb HindIII-EcoRI fragment of pHRGFPGUS	This work	promoter evaluation
pHREYFP	HindIII-EcoRI fragment from pEYFP (containing the *eyfp* gene) joined with the 5.8 kb HindIII-EcoRI fragment of pHRGFPGUS	This work	promoter evaluation

described in Sambrook and Russell (2001) [36]. The total DNA extraction of *A. amazonense* was performed as described by Wilson (1997) [37]. The primers used for PCR are listed in Table 2. All of the restriction and modification enzymes utilized in this work were purchased from New England Biolabs. The Taq DNA polymerase was provided by CenBiot Enzimas (Centro de Biotecnologia, UFRGS).

Isolation of *glnB* and *glnK* genes from *A. amazonense*

The genomic library enriched with 2-3 kb SalI DNA fragments was constructed as follows: the genomic DNA was digested with SalI and subsequently separated in agarose gel by electrophoresis. The 2-3 kb fragments were excised from the agarose gel and purified. Finally, these fragments were cloned in the pUC18 plasmid. This genomic library was partially sequenced and the *glnK* gene was identified using BLAST searches.

Inverse PCR for *glnB* isolation was performed according to Sambrook and Russell (2001) [36]. *Azospirillum amazonense* genomic DNA was digested with SalI and subsequently circularized. The PCR was performed with the glnB_sfint and revsf_glnBint primers (Table 2) and the circularized SalI DNA as a template. The 5' portion of the *glnA* gene was isolated by PCR with the revsf_glnBint and glnA_aa_do primers (Table 2) and the genomic DNA as a template.

The DNA sequencing was performed using a MEGABACE automated platform (Centro de Biotecnologia, UFRGS). Sequences were assembled using the Staden software package [39]. Gene annotation was carried out by Artemis software version 12.0 [40] along with BLAST software using the NCBI database http://blast.ncbi.nlm.nih.gov/Blast.cgi. Both sequences were deposited in the NCBI nucleotide database under the following access numbers: *glnB* region [GenBank:HM161849] and *glnK* region [GenBank:HM161850].

Antibiotic minimum inhibitory concentration test

The minimum inhibitory concentration of *A. amazonense* to the antibiotics (gentamicin, kanamycin, tetracycline, and ampicillin) was basically evaluated as described in Andrews (2001) [41]. The antibiotics were serially diluted in 1 mL of M79 medium at concentrations from 256 μg/mL to 0.5 μg/mL. An overnight culture of *A. amazonense* was diluted to 4×10^4 cells/mL. One milliliter of this dilution was added to one milliliter of M79 medium containing the appropriate antibiotic

Table 2 Primers utilized in this work

Primers	Sequence	Annealing temperature	Amplicon length (bp)	Purpose
glnB_sfint	CGCCGCGATACAGCTCGGTATG	57°C	2108	*glnB* region isolation
revsf_glnBint	GATGGACGATCAGTTGGTCGA	57°C	2108	*glnB* region isolation
glnA_aa_do	ACGGTCGGCACTTCCTTCAG	53°C	1621	*glnB* region isolation
pglnB_up_BamHI	CGGGATCCTCGTCGAAGCTGAAGGTCAT	55°C	727	*glnB* promoter amplification
pglnB_do_NcoI	GATCTTTTCCATGGCTTACGGC	55°C	727	*glnB* promoter amplification
pglnK_up_BamHI	CGGGATCCTTGTCCAGTGCCACGCTCAT	55°C	328	*glnK* promoter amplification
pglnK_do_NcoI	CACGAGCTCCATGGGTAGTCC	55°C	328	*glnK* promoter amplification
KglndeI_A _EcoRI	ATGAATTCCAATGCACAGGGTGCGTA	55°C	574 (AB) and 1111 (AD)	*glnK* mutagenesis
KglndeI_B	**CCCATCCACTAAACTTAAACA**CGCCACCACGAGCTTCAT	55°C	574	*glnK* mutagenesis
KglndeI_C	**TGTTTAAGTTTAGTGGATGGG**ATGACCATCGCCGACGCG	55°C	558	*glnK* mutagenesis
KglndeI_D_BamHI	CGGGATCCCGATGGTGGGCGGATATTTG	55°C	558 (CD) and 1111 (AD)	*glnK* mutagenesis
glnK_NdeI_up	GGACTACATATGAAGCTCGTGGTG	60°C	361 (wt) or 121 (mut)	*glnK* mutagenesis verification
glnK_BamHI_do	CGTCACGGGATCCTCATAAGGC	60°C	361 (wt) or 121 (mut)	*glnK* mutagenesis verification
conf_glnK_up	GCCCCCTCCAGGATCTTC	55°C	1522 (wt) or 1282 (mut)	*glnK* mutagenesis verification
conf_glnK_do	GGGTAAAATGCCCTTGTCCA	55°C	1522 (wt) or 1282 (mut)	*glnK* mutagenesis verification

Underline - restriction sites; Bold - sequence tag; wt - wild-type; mut - mutant; AB - amplification using the primers KglndeIA_EcoRI and KglndeIB; AD - amplification using the primers KglndeI_A_EcoRI and KglndeI_D_BamHI; CD - amplification using the primers KglndeIC and KglndeI_D_BamHI

concentration. The cells were cultivated in a rotary shaker at 150 rpm for 40 h at 35°C.

Conjugation

Conjugation was basically carried out as described by Clerico et al. (2007) [42]. However, some modifications were made as follows: overnight cultures of *A. amazonense* Y2 (receptor), *E. coli* XL1-Blue containing the plasmid pRK2013 (helper), and *E. coli* XL1-Blue containing the appropriate plasmid (donor) were used. Approximately 1 mL of the *A. amazonense* culture with an OD_{600} = 2 (1.3 × 10^9 cfu/ml) was mixed with 1 mL of each helper and donor cultures with an OD_{600} = 0.2 (2 × 10^8 cfu/mL) (ratio 10:1:1), unless stated otherwise. This mixture was harvested by centrifugation at 6000 *g* for 2 min and then resuspended in 100 μL of MLB medium (LB and M79 mixture at a proportion of 8:2), and this volume was then spotted onto MLB agar and incubated for 20 h at 35°C. Following this, the cell mass was resuspended in 200 μL of M79 medium and plated on M79 medium containing the appropriate antibiotic.

Electroporation

The preparation of cells was based on the protocol described by Schultheiss and Schüler (2003) [27]. A 3 mL overnight culture of *A. amazonense* was inoculated in 250 mL of M79 and the cells were cultivated to an OD_{600} of ~0.12 (early-log growth phase), unless stated otherwise. From this point, all manipulations were conducted on ice. The cells were incubated in ice for 30 min and then harvested by centrifugation at 5000 *g* for 20 min at 10°C. The cells were resuspended in 100 mL of electroporation buffer (pH 6.5 HEPES 1 mM, $MgCl_2$ 1 mM, and sucrose 200 mM) and again harvested by centrifugation (20 min at 5000 *g*). Subsequently, the cells were resuspended in 40 mL of electroporation buffer and again harvested by centrifugation. At the end, the cells were resuspended in 250 μL of electroporation buffer (final concentration of ~10^{10} cfu/mL), distributed in aliquots of 40 μL, and frozen in liquid nitrogen. Cell electroporation was carried out as follows: the 40 μL aliquot was mixed with 50 ng of the pHRGFPGUS vector and electroporated through a Gene Pulser apparatus (Bio-Rad Laboratories Inc.) with 12.5 kV/cm, 25 μF and 200 Ω, unless stated otherwise. After electrical discharge, the cells were resuspended in 500 μL of M79 medium and incubated at 35°C for 3 h in a rotary shaker at 150 rpm. Subsequently, the cells were plated on solid M79 medium containing 20 μg/mL of kanamycin and incubated for 2 days at 35°C.

Gene mutagenesis

Site-directed mutagenesis was based on a protocol described by Eggeling and Reyes (2005) [43]. In summary, the flanking regions of the *glnK* gene were amplified using the primers KglndelA_EcoRI/KglndelB and KglndelC/KglndelD_BamHI (Table 2). These amplification products were joined by Crossover PCR [31] using the primers KglndelA_EcoRI/KglndelD_BamHI (Table 2) and cloned in pK19MOBSACB digested with EcoRI and BamHI, generating the plasmid pKΔK (Table 1). Subsequently, the vector pKΔK was transferred to *A. amazonense* by conjugation, as previously described, except that the medium utilized was MLB containing maltose instead of sucrose (10 g/L) and ampicillin (100 μg/mL) for the counter-selection of *E. coli*. A kanamycin-resistant colony was isolated and cultured overnight in 3 mL of M79 (containing 10 g/L of maltose instead of sucrose). The culture was serially diluted and plated on M79 medium (containing 10 g/L of sucrose). Fifty sucrose-resistant colonies were replica plated onto both kanamycin-containing and pure M79 agar plates. Seven kanamycin-sensitive/sucrose-resistant colonies were submitted to Touchdown-PCR to identify those that had replaced the wild-type *glnK* gene with the mutant allele. The Touchdown-PCR was performed using the primers glnK_NdeI_up and glnK_BamHI_do (Table 2) under the following conditions: an initial denaturing step of 94°C for 5 min; 15 cycles of 94°C for 30 s, 60°C-56°C for 30 s (for each three cycles one degree was decreased), and 72°C for 30 s; 15 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s. The PCR utilizing the primers Conf_glnK_up and Conf_glnK_do (Table 2), which flank the recombination sites of the *glnK* region, was carried out in the same way as standard PCR procedures [36].

Gene reporter system

The upstream sequences of the genes utilized in this work were analyzed by Patser (available on the RSAT webserver) [44] with an *S. meliloti* sigma 70 factor weight matrix [33].

A series of reporter vectors was developed to evaluate the activity of different promoters (Table 1). The upstream regions of the *glnB* and *glnK* genes were amplified utilizing the primers listed in Table 2. Subsequently, these amplicons were cloned into the pEYFP vector at the NcoI and BamHI sites, generating pPBEYFP and pPKEYFP plasmids, respectively. After evaluation of the integrity of these amplicons by automated sequencing, the HindIII-EcoRI fragment, containing the promoter-*eyfp* fusion, was transferred to the HindIII-EcoRI fragment of pHRGFPGUS, which contains the replication origin, the mobilization site, and the kanamycin resistance marker, generating the pHRPBEYFP and pHRPKEYFP plasmids, respectively.

The pHRAATEYFP plasmid was constructed in the following way: the NcoI-BglII fragment of pAAGLNK, containing the upstream region of the *aat* gene, was

transferred to pEYFP, generating the plasmid pAATEYFP. The HindIII-EcoRI fragment from this plasmid was transferred to the HindIII-EcoRI fragment of pHRGFPGUS, generating pHRAATEYFP.

The negative control plasmid, which did not contain a promoter, was constructed as follows: the NcoI-BamHI fragment of pEYFP was transferred to the HindIII-EcoRI fragment of pHRGFPGUS, forming the plasmid pHREYFP.

The positive control plasmid pHRLACEYFP is a fusion of the major EcoRI-EcoRV fragment of pHRGFPGUS with the PvuII-EcoRI fragment of pEYFP.

All of the plasmids were transferred to *A. amazonense* by tri-parental mating or electroporation. The promoter activity assay was basically performed as described in MacLellan et al. (2006) [33]. *Azospirillum amazonense* containing the reporter vectors was cultivated in M79 medium overnight in a rotary shaker at 35°C. The cells were washed in sterile saline solution (0.85% NaCl) and resuspended in this same solution to an OD_{600} of between 0.06-0.39. Two hundred microlitres of the cell suspensions were deposited on black microtiter plates and fluorescence was measured with an excitation wavelength of 488 nm and an emission wavelength of 527 nm. The optical densities of the cell suspensions were measured at 600 nm on clear microtiter plates. Specific fluorescence was obtained by dividing the fluorescence by the optical density. Statistical analysis was performed using SAS JMP8 software: the specific fluorescence data was subjected to the natural logarithm to homogenize the variances (tested by Levene's test) and subsequently submitted for ANOVA/Tukey HSD tests ($P < 0.01$).

Acknowledgements and Funding

We especially thank Professor Emanuel E. Souza for kindly supplying the pHRGFPGUS plasmid. We thank Professor Marilene Henning Vainstein for kindly revising the manuscript. We also thank Professors Luciane Passaglia, Giancarlo Pasquali, Sídia Marques, and Carlos Termignoni for all of the assistance they provided. We also thank EMBRAPA-RJ for providing the *A. amazonense* Y2 strain.

This work was supported by grants from The Brazilian National Research Council (CNPq) and the Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS). FHS, DBT and SSW received scholarships from CAPES.

Author details

[1]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (Av. Bento Gonçalves, 9500, Campus do Vale), Porto Alegre, RS, Brazil. [2]Departamento de Biologia Molecular e Biotecnologia - Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (Av. Bento Gonçalves, 9500, Campus do Vale), RS, Brazil.

Authors' contributions

FHS conceived, coordinated and carried out the research study, drafted the manuscript, and created the illustrations and the tables. DSA performed the antibiotic minimum inhibitory concentration tests and helped with the electroporation procedures. DBT helped to isolate the *glnB* gene, designed some primers, and revised the manuscript. SSW helped with the reporter assays, and revised the manuscript. ISS conceived and coordinated the study, and revised the manuscript. All authors read and approved the final manuscript.

Received: 18 March 2011 Accepted: 16 May 2011
Published: 16 May 2011

References

1. Berg G: **Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture.** *Appl Microbiol Biotechnol* 2009, **84**:11-18.
2. Spiertz JHJ: **Nitrogen, sustainable agriculture and food security. A review.** *Agronomy for Sustainable Development* 2010, **30**:43-55.
3. Lucy M, Reed E, Glick BR: **Applications of free living plant growth-promoting rhizobacteria.** *Antonie Van Leeuwenhoek* 2004, **86**:1-25.
4. Bashan Y, De-Bashan L: **How the Plant Growth-Promoting Bacterium *Azospirillum* Promotes Plant Growth - A Critical Assessment.** *Adv agron* 2010, **108**:77-136.
5. Magalhães FMM, Baldani JI, Souto SM, Kuykendall JR, Döbereiner J: **A new acid-tolerant *Azospirillum* species.** *An Acad Bras Ciênc* 1983, **55**:417-430.
6. Baldani JI, Baldani VL: **History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience.** *An Acad Bras Ciênc* 2005, **77**:549-579.
7. Hartmann A, Fu H, Burris RH: **Regulation of nitrogenase activity by ammonium chloride in *Azospirillum* spp.** *J Bacteriol* 1986, **165**:864-870.
8. Rodrigues E, Rodrigues L, de Oliveira A, Baldani V, Teixeira KRS, Urquiaga S, Reis V: ***Azospirillum amazonense* inoculation: effects on growth, yield and N2 fixation of rice (*Oryza sativa* L.).** *Plant Soil* 2008, **302**:249-261.
9. Forchhammer K: **PII signal transducers: novel functional and structural insights.** *Trends Microbiol* 2008, **16**:65-72.
10. Leigh J, Dodsworth J: **Nitrogen regulation in bacteria and archaea.** *Annu Rev Microbiol* 2007, **61**:349-377.
11. Sant'Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC, Schrank IS: **The PII superfamily revised: a novel group and evolutionary insights.** *J Mol Evol* 2009, **68**:322-336.
12. Arcondéguy T, Jack R, Merrick M: **P(II) signal transduction proteins, pivotal players in microbial nitrogen control.** *Microbiol Mol Biol Rev* 2001, **65**:80-105.
13. Conroy M, Durand A, Lupo D, Li XD, Bullough P, Winkler F, Merrick M: **The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel.** *Proc Natl Acad Sci USA* 2007, **104**:1213-1218.
14. de Zamaroczy M: **Structural homologues P(II) and P(Z) of *Azospirillum brasilense* provide intracellular signalling for selective regulation of various nitrogen-dependent functions.** *Mol Microbiol* 1998, **29**:449-463.
15. Huergo L, Merrick M, Monteiro R, Chubatsu L, Steffens M, Pedrosa FO, Souza E: **In vitro interactions between the PII proteins and the nitrogenase regulatory enzymes dinitrogenase reductase ADP-ribosyltransferase (DraT) and dinitrogenase reductase-activating glycohydrolase (DraG) in *Azospirillum brasilense*.** *J Biol Chem* 2009, **284**:6674-6682.
16. Araújo L, Monteiro R, Souza E, Steffens B, Rigo L, Pedrosa FO, Chubatsu L: **GlnB is specifically required for *Azospirillum brasilense* NifA activity in *Escherichia coli*.** *Res Microbiol* 2004, **155**:491-495.
17. Chen S, Liu L, Zhou X, Elmerich C, Li JL: **Functional analysis of the GAF domain of NifA in *Azospirillum brasilense*: effects of Tyr->Phe mutations on NifA and its interaction with GlnB.** *Mol Genet Genomics* 2005, **273**:415-422.
18. Fu HA, Hartmann A, Lowery RG, Fitzmaurice WP, Roberts GP, Burris RH: **Posttranslational regulatory system for nitrogenase activity in *Azospirillum* spp.** *J Bacteriol* 1989, **171**:4679-4685.
19. Davison J: **Genetic tools for pseudomonads, rhizobia, and other gram-negative bacteria.** *BioTechniques* 2002, **32**:386-388.
20. Schweizer H: **Bacterial genetics: past achievements, present state of the field, and future challenges.** *BioTechniques* 2008, **44**:633-641.
21. Holguin G, Patten CL, Glick BR: **Genetics and molecular biology of *Azospirillum*.** *Biol Fertil Soils* 1999, **29**:10-23.
22. Aune T, Aachmann F: **Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed.** *Appl Microbiol Biotechnol* 2010, **85**:1301-1313.
23. Lurquin PF: **Gene transfer by electroporation.** *Mol Biotechnol* 1997, **7**:5-35.
24. Taketo A: **Electrotransformation of bacteria.** In *Electromanipulation of Cells.* Edited by: Zimmermann U, Neil GA. Boca Raton, FL: CRC Press; 1996:107-136.

25. Vande Broek A, Gool A, Vanderleyden J: **Electroporation of *Azospirillum brasilense* with plasmid DNA.** *FEMS Microbiol Lett* 1989, **61**:177-182.
26. Wirth R, Friesenegger A, Fiedler S: **Transformation of various species of gram-negative bacteria belonging to 11 different genera by electroporation.** *Mol Genet Genomics* 1989, **216**:175-177.
27. Schultheiss D, Schüler D: **Development of a genetic system for *Magnetospirillum gryphiswaldense*.** *Arch Microbiol* 2003, **179**:89-94.
28. Lerner A, Castro-Sowinski S, Valverde A, Lerner H, Dror R, Okon Y, Burdman S: **The *Azospirillum brasilense* Sp7 *noeJ* and *noeL* genes are involved in extracellular polysaccharide biosynthesis.** *Microbiology* 2009, **155**:4058-4068.
29. Lerner A, Castro-Sowinski S, Lerner H, Okon Y, Burdman S: **Glycogen phosphorylase is involved in stress endurance and biofilm formation in *Azospirillum brasilense* Sp7.** *FEMS Microbiol Lett* 2009, **300**:75-82.
30. Xie Z, Ulrich L, Zhulin I, Alexandre G: **PAS domain containing chemoreceptor couples dynamic changes in metabolism with chemotaxis.** *Proc Natl Acad Sci USA* 2010, **107**:2235-2240.
31. Link AJ, Phillips D, Church GM: **Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization.** *J Bacteriol* 1997, **179**:6228-6237.
32. Gourse R, Ross W, Rutherford S: **General Pathway for Turning on Promoters Transcribed by RNA Polymerases Containing Alternative sigma Factors.** *J Bacteriol* 2006, **188**:4589-4591.
33. MacLellan S, MacLean A, Finan T: **Promoter prediction in the rhizobia.** *Microbiology* 2006, **152**:1751-1763.
34. Holguin G, Glick BR: **Expression of the ACC Deaminase Gene from *Enterobacter cloacae* UW4 in *Azospirillum brasilense*.** *Microb Ecol* 2001, **41**:281-288.
35. Fred EB, Waskman SA: *Laboratory manual of general microbiology with special reference to the microorganisms of the soil* New York: McGraw-Hill Book Company, Inc; 1928.
36. Sambrook J, Russell DW: *Molecular Cloning: A Laboratory Manual* New York: Cold Spring Harbor Laboratory; 2001.
37. Wilson K: **Preparation of genomic DNA from bacteria.** In *Current protocols in molecular biology*.. 1 edition. Edited by: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al. New York: Wiley; 1997:2.
38. Potrich DP, Passaglia LM, Schrank IS: **Partial characterization of nif genes from the bacterium *Azospirillum amazonense*.** *Braz J Med Biol Res* 2001, **34**:1105-1113.
39. Staden R, Beal KF, Bonfield JK: **The STADEN package.** *Methods Mol Biol* 1998, **132**:115-130.
40. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: **Artemis: sequence visualization and annotation.** *Bioinformatics* 2000, **16**:944-945.
41. Andrews J: **Determination of minimum inhibitory concentrations.** *J Antimicrob Chemother* 2001, **48**:5-16.
42. Clerico EM, Ditty JL, Golden SS: **Specialized techniques for site-directed mutagenesis in cyanobacteria.** *Methods Mol Biol* 2007, **362**:155-171.
43. Eggeling L, Reyes O: **Deletion of chromosomal sequences and allelic exchange.** In *Handbook of Corynebacterium glutamicum*. Edited by: Eggeling L, Bott M. Boca Raton: CRC press; 2005:557-559.
44. Thomas-Chollier M, Sand O, Turatsinze JV, Janky R, Defrance M, Vervisch E, Brohée S, van Helden J: **RSAT: regulatory sequence analysis tools.** *Nucleic Acids Res* 2008, **36**:W119-W127.
45. Figurski D, Helinski D: **Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans.** *Proc Natl Acad Sci USA* 1979, **76**:1648-1652.
46. Ramos HJ, Roncato-Maccari LD, Souza EM, Soares-Ramos JR, Hungria M, Pedrosa FO: **Monitoring *Azospirillum*-wheat interactions using the *gfp* and *gusA* genes constitutively expressed from a new broad-host range vector.** *J Biotechnol* 2002, **97**:243-252.
47. Covert SF, Kapoor P, Lee MH, Briley A, Nairn CJ: ***Agrobacterium tumefaciens*-mediated transformation of *Fusarium circinatum*.** *Mycol Res* 2001, **105**:259-264.
48. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A: **Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*.** *Gene* 1994, **145**:69-73.

doi:10.1186/1471-2180-11-107
Cite this article as: Sant'Anna *et al.*: Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*. *BMC Microbiology* 2011 **11**:107.

Supplementary material

Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*

Sant'Anna, FH[1]; Andrade, DS[1]; Trentini, DB[1]; Weber, SS[1]; Schrank, IS[2§]

1- Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (Av. Bento Gonçalves, 9500, Campus do Vale), Porto Alegre, RS, Brazil

2- Departamento de Biologia Molecular e Biotecnologia - Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (Av. Bento Gonçalves, 9500, Campus do Vale), RS, Brazil

[§]Corresponding author

E-mail addresses:

FHS (fhsantanna@yahoo.com.br)

DSA (dieime.sa@gmail.com)

DBT (deborabtrentini@gmail.com)

SSW (shana@cbiot.ufrgs.br)

ISS (irene@cbiot.ufrgs.br)

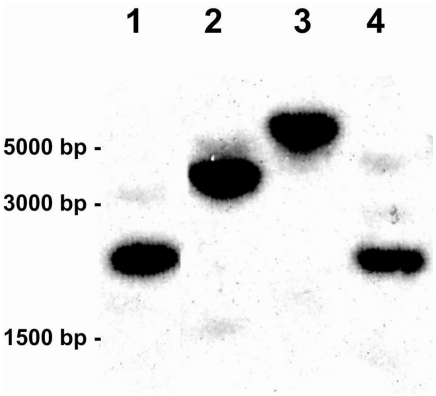


Figure S1 - Identification of PII gene homologs in the *A. amazonense* genome by Southern Blot

Hybridization pattern utilizing a radioactive *glnB* probe over *A. amazonense* total DNA digested with different restriction enzymes: 1 - SalI, 2 - NcoI, 3 - PstI, and 4 - HincII

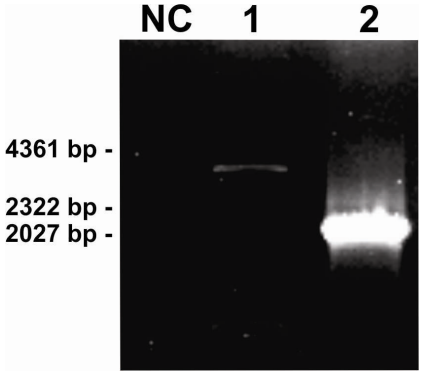


Figure S2 - Isolation of the *glnB* region by Inverse PCR.

Agarose gel electrophoresis of Inverse PCR amplification products

NC - negative control, 1 - NcoI circularized DNA, 2 - SalI circularized DNA

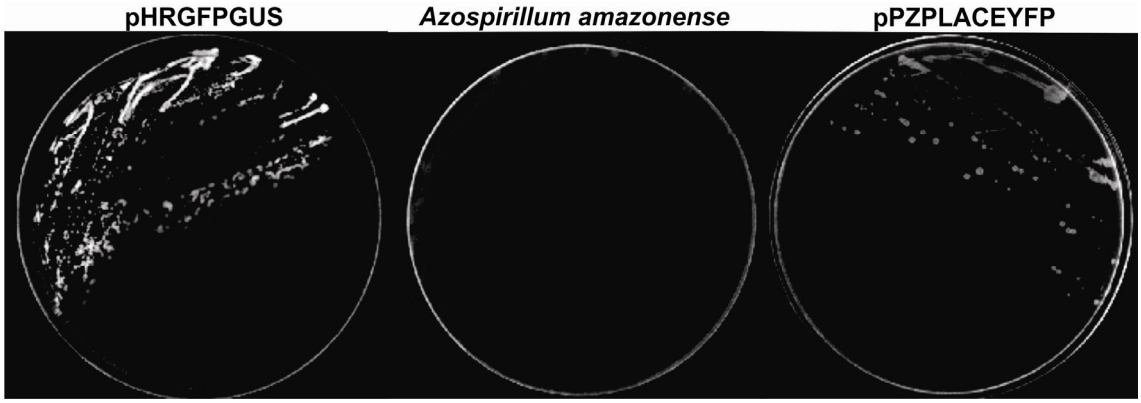


Figure S3 - Fluorescence detection of the *A. amazonense* transconjugants

Wild-type *A. amazonense* is in the middle. *A. amazonense* bearing the pHRGFPGUS plasmid is on the left. *A. amazonense* bearing the pPZPLACEYFP plasmid is on the right.

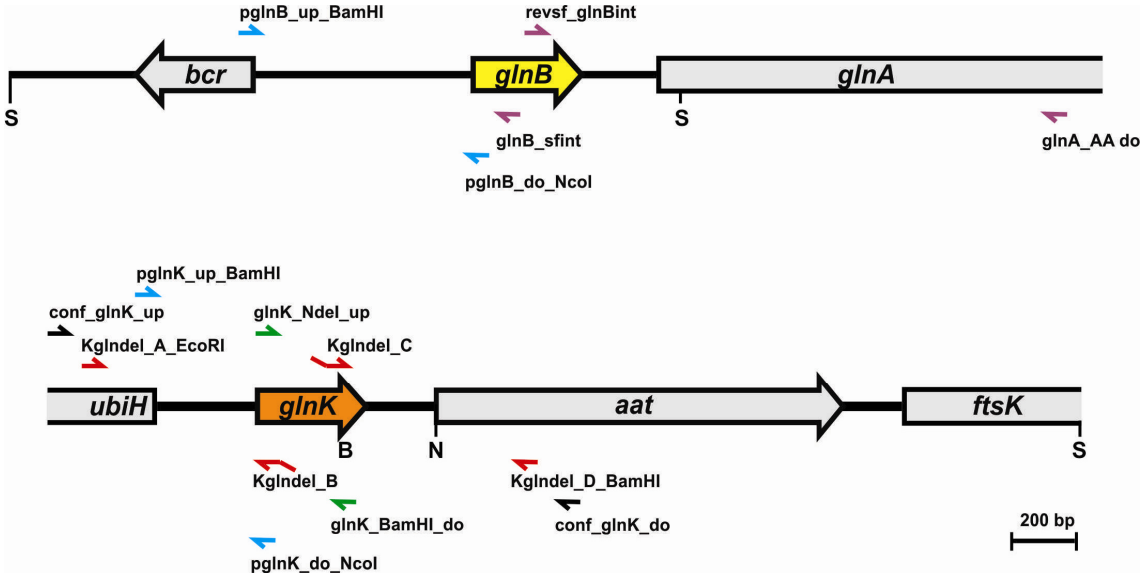
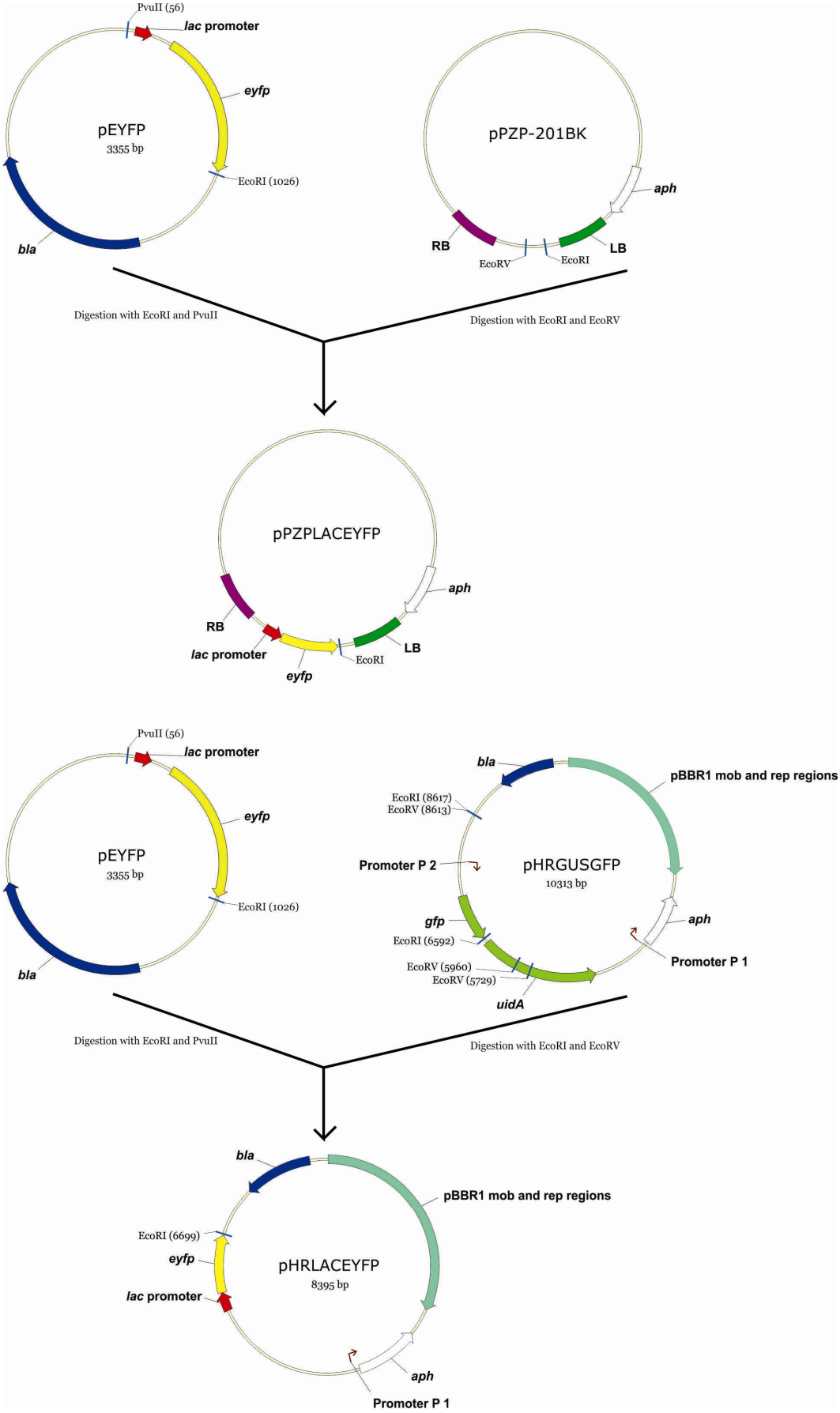
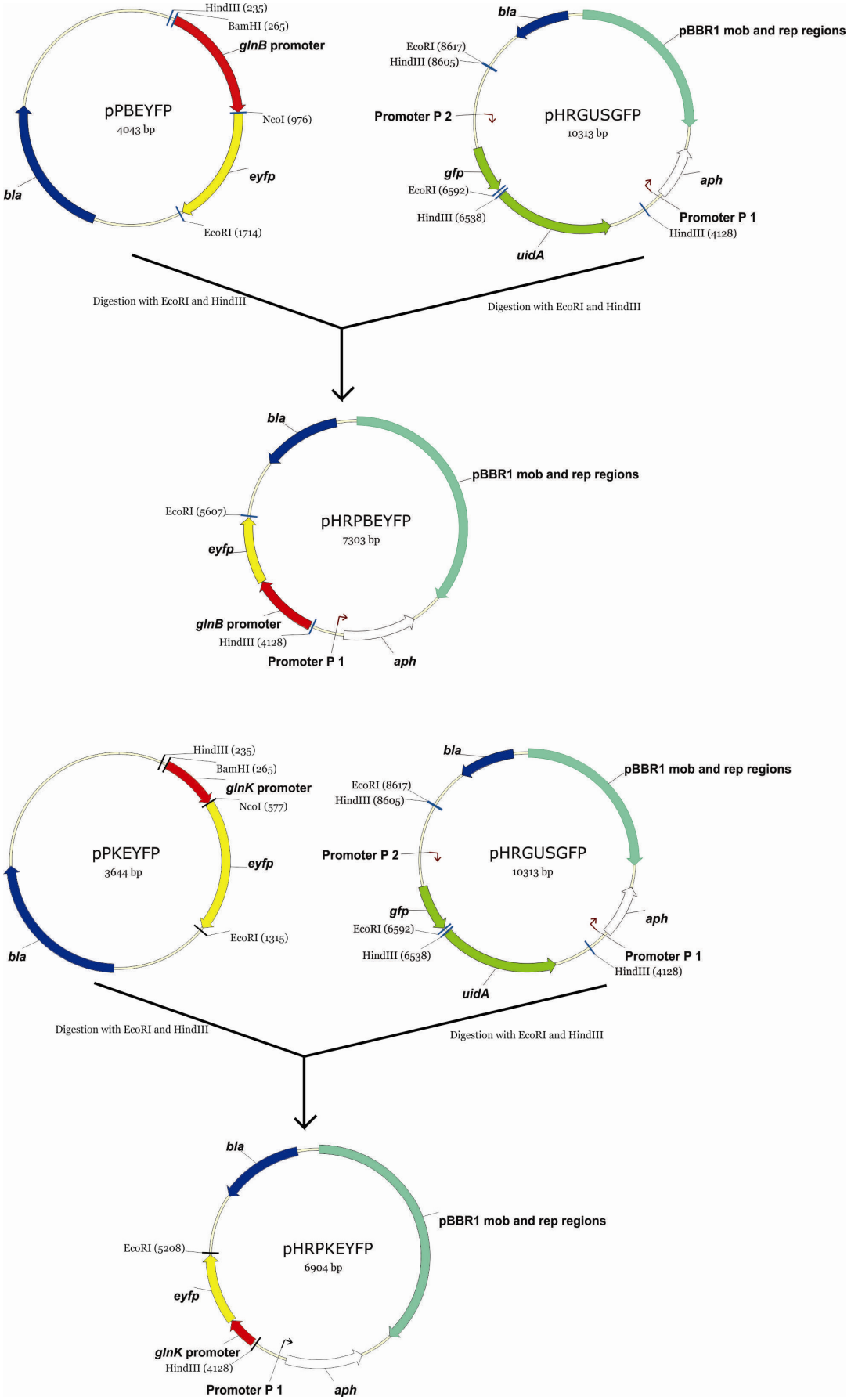


Figure S4 - Physical maps showing the primer binding sites at the *glnK* and *glnB* regions of *A. amazonense*

Genes are represented by the large arrows. The small arrows represent the primer binding sites. Restriction sites are indicated by capital letters: S- SalI, B - BglII and N- NcoI.


PvuII (56)
lac promoter
eyfp
pEYFP
3355 bp
EcoRI (1026)
bla
pPZP-201BK
aph
RB
LB
EcoRV
EcoRI
Digestion with EcoRI and PvuII
Digestion with EcoRI and EcoRV
pPZPLACEYFP
aph
RB
LB
lac promoter
EcoRI
eyfp
PvuII (56)
lac promoter
eyfp
pEYFP
3355 bp
EcoRI (1026)
bla
bla
pBBR1 mob and rep regions
EcoRI (8617)
EcoRV (8613)
Promoter P 2
pHRGUSGFP
10313 bp
aph
gfp
EcoRI (6592)
Promoter P 1
EcoRV (5960)
EcoRV (5729)
uidA
Digestion with EcoRI and PvuII
Digestion with EcoRI and EcoRV
bla
pBBR1 mob and rep regions
EcoRI (6699)
eyfp
pHRLACEYFP
8395 bp
lac promoter
aph
Promoter P 1


HindIII (235)
BamHI (265)
glnB promoter
NcoI (976)
pPBEYFP
4043 bp
eyfp
bla
EcoRI (1714)
bla
pBBR1 mob and rep regions
EcoRI (8617)
HindIII (8605)
Promoter P 2
pHRGUSGFP
10313 bp
gfp
EcoRI (6592)
HindIII (6538)
aph
Promoter P 1
HindIII (4128)
uidA
Digestion with EcoRI and HindIII
Digestion with EcoRI and HindIII
bla
pBBR1 mob and rep regions
EcoRI (5607)
pHRPBEYFP
7303 bp
eyfp
glnB promoter
HindIII (4128)
Promoter P 1
aph
HindIII (235)
BamHI (265)
glnK promoter
NcoI (577)
pPKEYFP
3644 bp
eyfp
bla
EcoRI (1315)
bla
pBBR1 mob and rep regions
EcoRI (8617)
HindIII (8605)
Promoter P 2
pHRGUSGFP
10313 bp
gfp
EcoRI (6592)
HindIII (6538)
aph
Promoter P 1
HindIII (4128)
uidA
Digestion with EcoRI and HindIII
Digestion with EcoRI and HindIII
bla
pBBR1 mob and rep regions
EcoRI (5208)
pHRPKEYFP
6904 bp
eyfp
glnK promoter
HindIII (4128)
Promoter P 1
aph

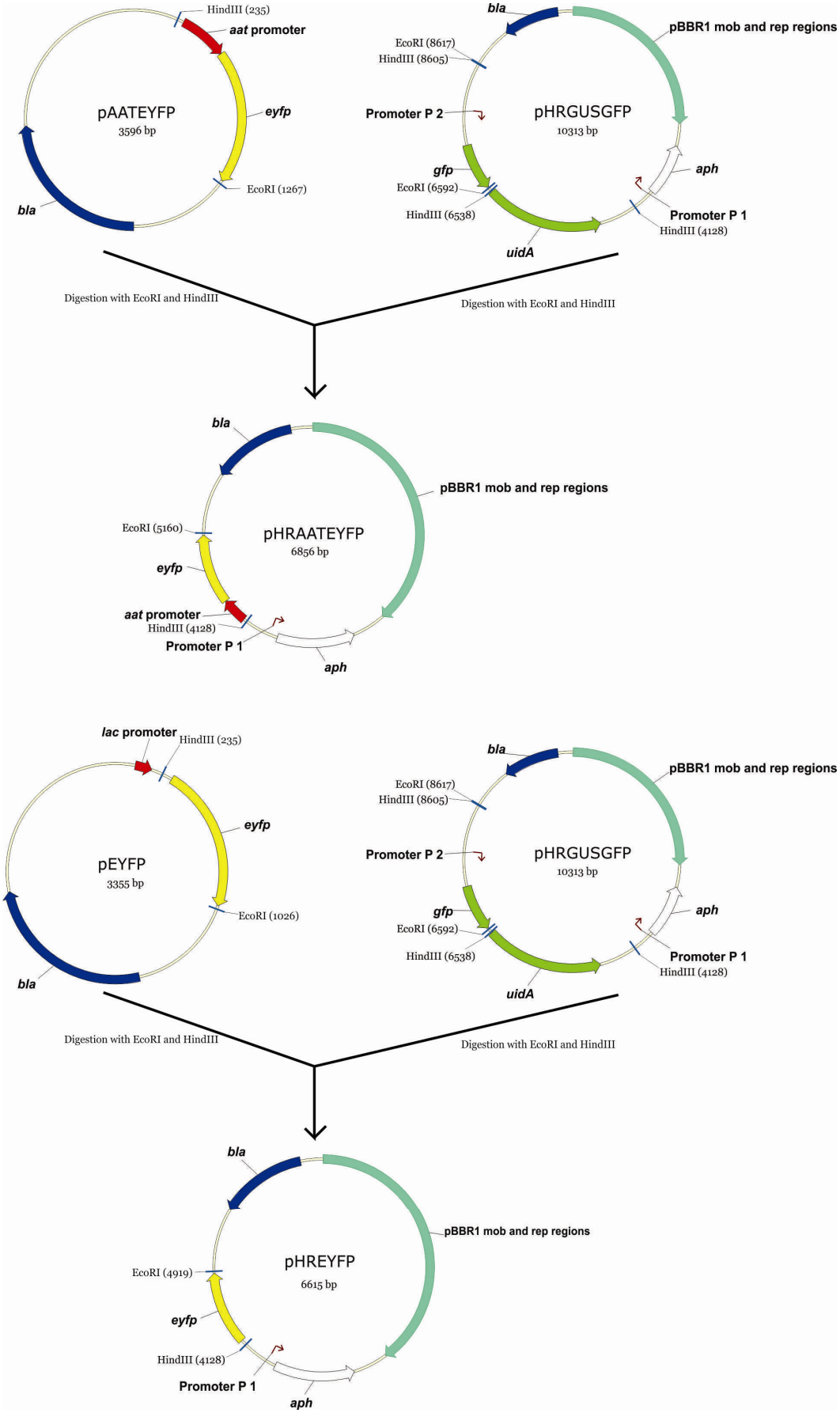

HindIII (235)
aat promoter
pAATEYFP
3596 bp
eyfp
EcoRI (1267)
bla
bla
pBBR1 mob and rep regions
EcoRI (8617)
HindIII (8605)
Promoter P 2
pHRGUSGFP
10313 bp
gfp
EcoRI (6592)
HindIII (6538)
aph
Promoter P 1
HindIII (4128)
uidA
Digestion with EcoRI and HindIII
Digestion with EcoRI and HindIII
bla
pBBR1 mob and rep regions
EcoRI (5160)
pHRAATEYFP
6856 bp
eyfp
aat promoter
HindIII (4128)
Promoter P 1
aph
lac promoter
HindIII (235)
eyfp
pEYFP
3355 bp
EcoRI (1026)
bla
bla
pBBR1 mob and rep regions
EcoRI (8617)
HindIII (8605)
Promoter P 2
pHRGUSGFP
10313 bp
gfp
EcoRI (6592)
HindIII (6538)
aph
Promoter P 1
HindIII (4128)
uidA
Digestion with EcoRI and HindIII
Digestion with EcoRI and HindIII
bla
pBBR1 mob and rep regions
EcoRI (4919)
pHREYFP
6615 bp
eyfp
HindIII (4128)
Promoter P 1
aph

Figure S5 - Schematic diagrams showing the construction of the reporter vectors

The constructions are depicted in Table 1. *aph* - kanamycin resistance cassette, *bla* – ampicillin resistance cassette, *eyfp* – enhanced yellow fluorescent protein gene, *gfp* – green fluorescent protein, LB – left border, mob – mobilization, RB – right border, rep – replication, *uidA* - beta-glucuronidase (GUS) gene.

CAPÍTULO 2

Artigo publicado na revista BMC Genomics (Fator de impacto: 4,21)

Genomic insights into the versatility of the plant growth-promoting bacterium *Azospirillum amazonense*

Fernando H Sant'Anna, Luiz GP Almeida, Ricardo Cecagno, Luciano A Reolon, Franciele M Siqueira, Maicon RS Machado, Ana TR Vasconcelos & Irene S Schrank

BMC Genomics. 2011; 12: 409

O seguinte trabalho foi resultado de um esforço em conjunto do nosso laboratório com o Laboratório Nacional de Computação Científica (LNCC, Petrópolis, RJ). Ao último laboratório coube o sequenciamento, montagem e anotação automática das sequências. O nosso laboratório foi responsável pela validação manual das ORFs e pela análise dos resultados. Participei de praticamente todas as etapas do trabalho, entretanto, intensivamente das partes que se referem ao metabolismo do nitrogênio e do carbono, *quorum sensing* e taxonomia.

BMC Genomics

RESEARCH ARTICLE Open Access

Genomic insights into the versatility of the plant growth-promoting bacterium *Azospirillum amazonense*

Fernando H Sant'Anna[1], Luiz GP Almeida[2], Ricardo Cecagno[1], Luciano A Reolon[1], Franciele M Siqueira[1], Maicon RS Machado[1], Ana TR Vasconcelos[2] and Irene S Schrank[1,3*]

Abstract

Background: The species *Azospirillum amazonense* belongs to a well-known genus of plant growth-promoting bacteria. This bacterium is found in association with several crops of economic importance; however, there is a lack of information on its physiology. In this work, we present a comprehensive analysis of the genomic features of this species.

Results: Genes of *A. amazonense* related to nitrogen/carbon metabolism, energy production, phytohormone production, transport, quorum sensing, antibiotic resistance, chemotaxis/motility and bacteriophytochrome biosynthesis were identified. Noteworthy genes were the nitrogen fixation genes and the nitrilase gene, which could be directly implicated in plant growth promotion, and the carbon fixation genes, which had previously been poorly investigated in this genus. One important finding was that some *A. amazonense* genes, like the nitrogenase genes and RubisCO genes, were closer phylogenetically to Rhizobiales members than to species of its own order.

Conclusion: The species *A. amazonense* presents a versatile repertoire of genes crucial for its plant-associated lifestyle.

Background

The genus *Azospirillum* (α-proteobacteria class) encompasses free-living bacteria that can improve the growth of many economically important plants, mainly cereals (for an extensive review, see [1]). Therefore, these microorganisms are considered as plant growth-promoting rhizobacteria (PGPR). Species of this genus are widely distributed in nature, living in soils of tropical, subtropical and temperate regions all over the world. Several aspects of their physiology seem to be related to a plant stimulatory effect, notably their ability to synthesize phytohormones. Although these microorganisms are able to fix atmospheric nitrogen, the exact contribution of this process to plant growth is still disputable [1-3].

So far, fifteen species of the *Azospirillum* genus have been described ([4] and references therein). However, most research efforts have been dedicated to the species *Azospirillum brasilense*, neglecting the potential offered by the biological diversity of this genus.

The bacterium *A. amazonense*, the focus species of this study, was initially isolated from forage grasses grown in the Amazon region. Further studies revealed its broad ecological distribution, as it is also found in association with the roots of gramineous plants like rice, maize, sugarcane and sorghum [5,6]. This species is phylogenetically closer to *Azospirillum irakense* and *Rhodospirillum centenum* (also known as *Rhodocista centenaria*) than to *A. brasilense*. Unlike the latter, *A. amazonense* can use sucrose as a sole carbon source and is better adapted to acid environments [6].

In order to access the valuable information that genomic sequences can provide on the physiology of azospirilla, there have been independent efforts by research groups in sequencing their genomes. Currently, the genomes of three members of the *Azospirillum-R. centenum* group are available: *Azospirillum* sp. B510 [7], *A. brasilense* Sp245 [8] and *R. centenum* SW [9]. Although

* Correspondence: irene@cbiot.ufrgs.br
[1]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Campus do Vale, Porto Alegre, RS, Brazil
Full list of author information is available at the end of the article

many years have passed since the discovery of *A. amazonense*, there is scarce information about this species. Nevertheless, a recent study under greenhouse conditions showed that especially *A. amazonense* Y2 (wild-type strain) contributed to the growth of rice plants by means of biological nitrogen fixation [10], showing its potential for use as an agricultural inoculant. Therefore, the objective of our study was to sequence the *A. amazonense* Y2 genome and to analyze specific regions that could exert fundamental roles in its survival in the soil and in its ability to promote plant growth.

Results and Discussion

General features of the A. amazonense draft genome database

The *A. amazonense* Y2 presents four replicons with the following estimated sizes: 2.7 Mbp, 2.2 Mbp, 1.7 Mbp and 0.75 Mbp [11] The genomic G+C content of *A. amazonense* Y2 is 66.89%. The draft genome sequence consists of 7.044.835 bp divided in 1617 contigs. The average gene length is 1080. Currently, there are 3319 predicted CDS, where 2299 have sequence similarity to known genes, 501 are homologs to genes of unknown function and 519 are hypothetical genes exclusive to *A. amazonense*.

Taxonomic features of A. amazonense

In this study, a phylogenetic tree was constructed using the 16S rDNA sequences from microorganisms belonging to the orders Rhodospirillales and Rhizobiales (Figure 1). The resulting phylogenetic tree clearly shows a split between these orders. The outermost clade containing all the *Azospirillum* species divides in two main subclades: one containing *A. amazonense, A. irakense, Rhodocista pekingensis* and *R. centenum*, and another containing the other *Azospirillum* species. This result is in agreement with previous studies, showing the close evolutionary relationship between *A. amazonense* and *R. centenum* [12-14], and is also supported by the greater number of *A. amazonense* genes (22%) with best BLAST hits (KEGG Database) to *R. centenum* genes.

Another relevant feature of the 16S rDNA phylogenetic reconstruction is that *R. centenum* does not cluster with other *Rhodospirillum* species, suggesting that the genus nomenclature of *R. centenum* is not appropriate, as has been pointed out by previous publications [13,15].

Carbon metabolism

Azospirilla display versatile carbon metabolism in order to support their lives in the soil. *A. amazonense* is capable of growing on various disaccharides, hexoses and pentoses and a previous study suggested that *A. amazonense* is able to catabolize carbohydrates exclusively through the Entner-Doudoroff pathway (ED pathway) [16]. The genes encoding the key enzymes of this pathway, 6-phosphogluconate dehydratase and 2-dehydro-3-deoxy-phosphogluconate aldolase (KDPG aldolase) (Additional file 1), were found in the *A. amazonense* genome and seem to be organized as an operon. This same study also suggested that the glycolysis pathway (Embden-Meyerhof-Parnas pathway) was inoperative in *A. amazonense*, because no activity of 6-phosphofructokinase and fructose bisphosphate aldolase was detected in crude extracts [16]. However, predicted genes encoding those enzymes were found in the *A. amazonense* genome (Additional file 1). Therefore, although the genomic approach indicates that most probably *A. amazonense* is also able to consume carbohydrates via glycolysis, this catabolic feature should be experimentally retested.

As stated previously, one of the main differences between *A. amazonense* and *A. brasilense* is that *A. amazonense* is capable of consuming sucrose as the sole carbon source [6]. In the genome of *A. amazonense*, a predicted gene that codes for a putative α-glucosidase was identified (Additional file 1). This enzyme converts sucrose to glucose and fructose, substrates that can be promptly consumed by catabolic pathways.

The *A. amazonense* genome also harbors homologs of the genes *salB* and *salA* of *A. irakense* (Additional file 1). These genes encode β-glucosidases, enzymes implicated in the acquisition of glucose by means of the hydrolysis of aryl-β-glucosides, such as salicin [17].

Bacteria of the genus *Azospirillum* produce high levels of poly-β-hydroxybutyrate (PHB), the energy and carbon storage source utilized under nutritional stress conditions [2,18]. The essential genes for PHB biosynthesis are present in the *A. amazonense* genome: *phbA* (β-ketothiolase), *phbB* (aceto acetyl coenzyme A reductase) and *phbC* (PHB synthase) (Additional file 1). Furthermore, the *phaZ* gene that encodes a PHB depolymerase (Additional file 1), the first enzyme of the PHB degradation pathway, was also found in its genome.

One of the most surprising features of the *A. amazonense* genome is the presence of a gene cluster implicated in carbon fixation (the Calvin-Benson-Basham cycle) (Figure 2 and Additional file 1). The main genes of this cluster are the genes *cbbL* and *cbbS*, and they encode, respectively, the large and small subunits of ribulose-1,5-bisphosphate carboxylase (RubisCO). This enzyme is responsible for the incorporation of carbon dioxide in a molecule of ribulose-1,5-bisphosphate, generating two molecules of 3-phosphoglycerate, which can subsequently be used in biosynthetic pathways. A phylogenetic analysis of the concatenated RubisCO small and large subunits of *A. amazonense* revealed that they belong to the Form IC of RubisCOs (Figure 3). This type of enzyme is commonly found in α-

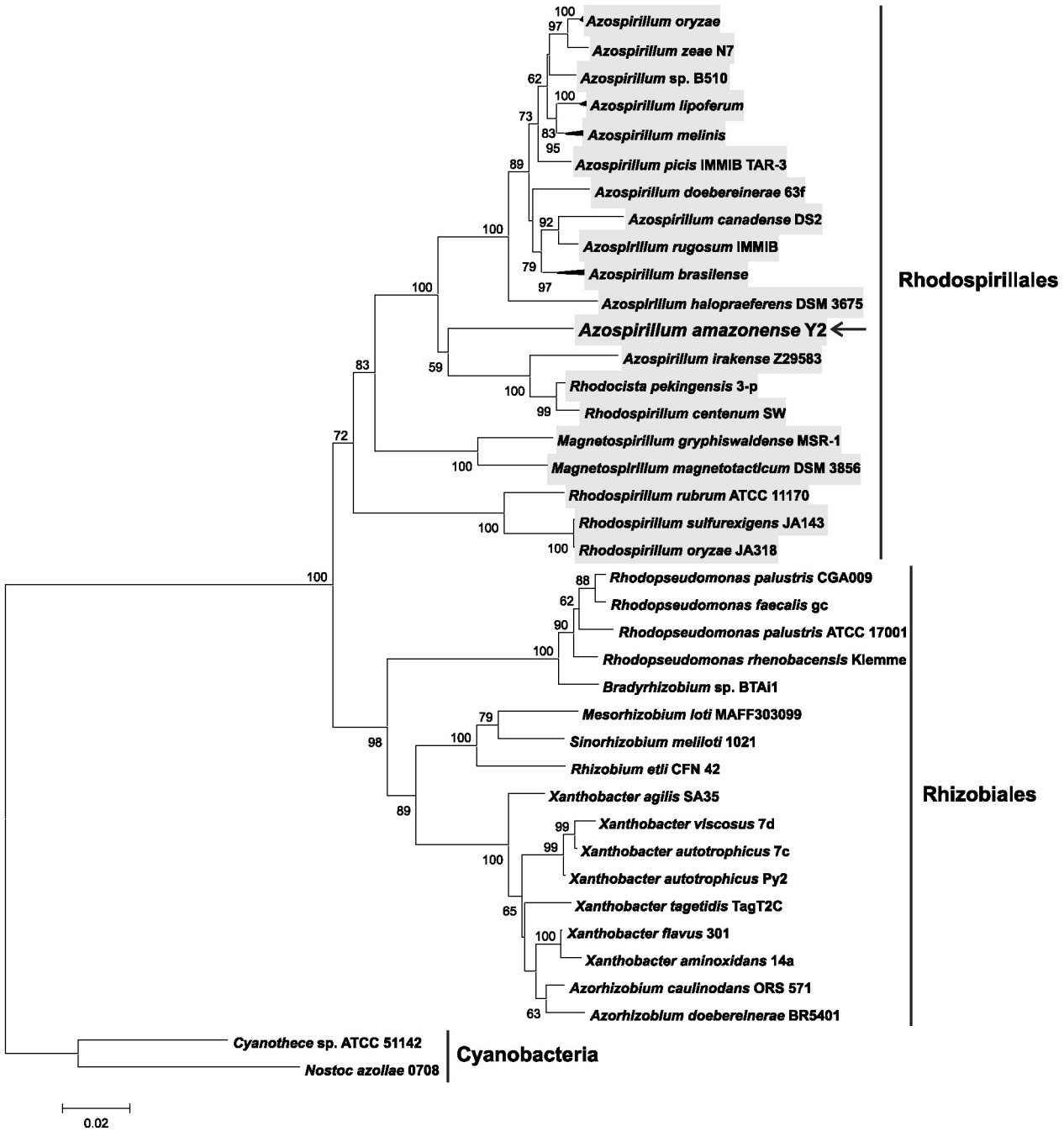


Figure 1 The evolutionary history of 16S rDNA sequences from members of the orders Rhodospirillales and Rhizobiales. The tree was constructed using the Neighbor-Joining method with 1000 bootstrap replicates (percentages greater than 50% are shown next to the branches). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. There were a total of 1254 positions in the final dataset. The species *Azospirillum amazonense* is indicated by an arrow. The gray boxes indicate members of the order Rhodospirillales.

Proteobacteria and it is adapted to environments with medium to high CO_2 and the presence of O_2 (in general, RubisCOs also have affinity to O_2 and high levels of this molecule can inhibit CO_2 fixation) [19]. So far, there have been no reports showing that *A. amazonense* has autotrophic behavior. However, from the *Azospirillum* group, at least *R. centenum* and *A. lipoferum* are known to be capable of growing autotrophically by means of RubisCO [9,20], unlike *Azospirillum* sp. B510 and *A. brasilense* Sp245, which do not

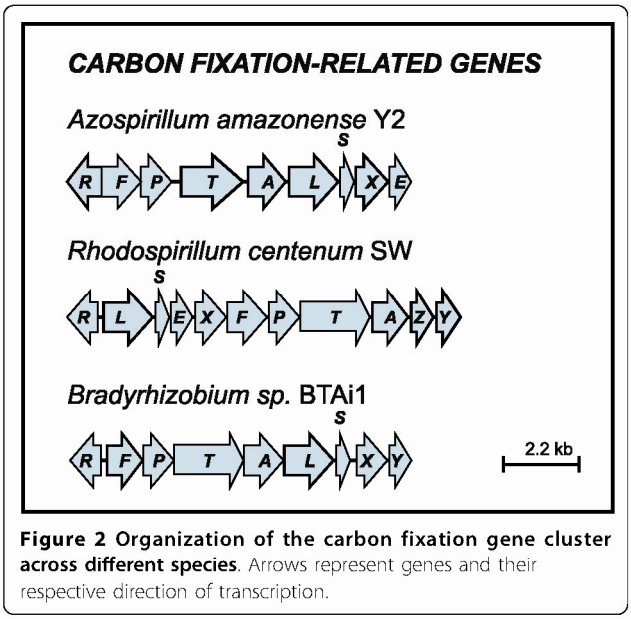


Figure 2 Organization of the carbon fixation gene cluster across different species. Arrows represent genes and their respective direction of transcription.

contain Form I or II of RubisCOs ("true" RubisCOs) encoded in their genomes.

The RubisCO phylogenetic reconstruction also indicated the close relationship of the *A. amazonense* enzyme with those from members of the family Bradyrhizobiaceae (order Rhizobiales) (Figure 3), namely *Rhodopseudomonas palustris* and *Bradyrhizobium* spp. In fact, the genetic organization of the carbon-fixation cluster of *A. amazonense* resembles that found in *Bradyrhizobium* sp. BTAi1, contrasting with the genetic organization of *R. centenum*. These incongruities, i.e. the genetic organization and phylogenetic relationship closer to Bradyrhizobiaceae members than to *R. centenum*, suggests that horizontal gene transfer may be an important driving force in the evolution and dispersion of RubisCOs in Proteobacteria.

Nitrogen metabolism

Azospirillum species are able to utilize distinct nitrogen sources, including ammonia, nitrate, nitrite, dinitrogen and amino acids [3,21]. The *A. amazonense* species has several genes implicated in nitrogen metabolism, which encode transporters, enzymes and regulatory proteins (Additional file 1).

Ammonia is the central compound of nitrogen metabolism and the preferred nitrogen source of many microorganisms. In general, nitrogen sources other than ammonia are converted into it to be assimilated [22]. The *A. amazonense* genome contains genes that are implicated in this conversion of alternative nitrogen sources, like nitrate/nitrite, urea and dinitrogen (Additional file 1). Once available, ammonia can be incorporated into the metabolism by the glutamine synthetase (GS)/glutamine:oxoglutarate aminotransferase (GOGAT) pathway, the genes for which are also encoded in the *A. amazonense* genome (Additional file 1).

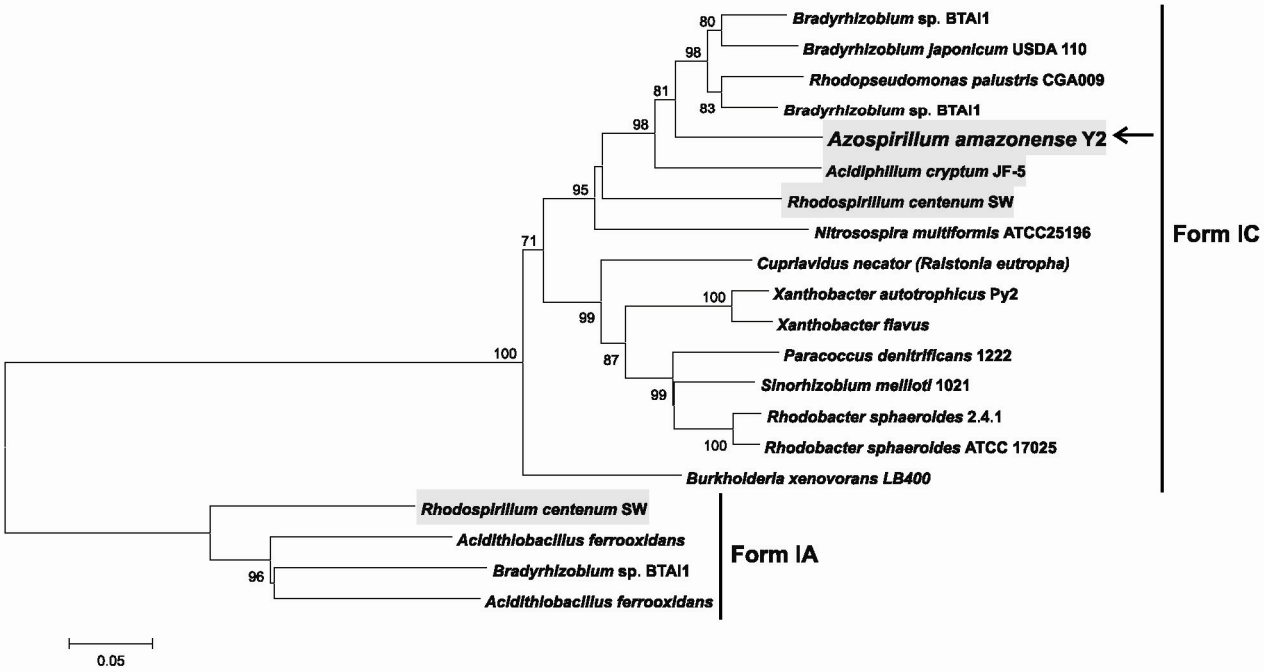


Figure 3 Phylogenetic tree based on the concatenated CbbL-CbbS RubisCO amino acid sequences. There were a total of 557 positions in the final dataset. Details are as shown in Figure 1.

Overall, the conversion of nitrogen compounds to ammonia expends some energy and, therefore, the metabolic pathways implicated in this process are strictly regulated to minimize energy waste. The central regulators of nitrogen metabolism are the PII proteins [22,23]. Three PII homolog genes (*glnB, glnK* and *glnK2*) were found in the *A. amazonense* genome (Additional file 1). The *glnK* gene and the *glnB* gene have ortholog counterparts in *Azospirillum* sp. B510, *A. brasilense* Sp245 and *R. centenum*. The *glnK* gene is upstream of the *aat* gene (aminotransferase) and the *glnB* is upstream of the *glnA* gene (glutamine synthetase) [24]. The third gene, *glnK2*, which is located downstream of the *amtB* gene, is absent in *A. brasilense* Sp245, *Azospirillum* sp. B510 and *R. centenum*, although this genetic association is frequently found in diverse prokaryotes [25].

The PII protein interactions with transporters, transcription factors and regulatory enzymes are well-established in the literature (for a review, see [22,23]) and the potential interaction targets found in the *A. amazonense* genome will be briefly discussed. One putative target is the *glnD* gene that codes for an uridilyl-transferase, an enzyme that uridylylates the PII proteins under nitrogen-limiting conditions [23]. Other potential targets found in the *A. amazonense* genome are two ammonium transporters, encoded by the *amtB* genes, which in the presence of high nitrogen levels are inhibited by PII proteins [26]. The adenylyltransferase enzyme (encoded by the *glnE* gene, Additional file 1) which regulates glutamine synthetase via covalent modifications [27] could also interact with PII proteins [28].

The analysis of the *A. amazonense* genome revealed the presence of the NtrBC and NifA systems (Additional file 1), which are PII-regulated signal transduction systems responsible for the coordination of genes implicated in the scavenging of alternative nitrogen sources [23]. Both NtrC and NifA rely on the presence of the sigma N factor (also known as RpoN, or sigma 54) to activate the transcription of specific genes [29], which is also present in the *A. amazonense* genome (Additional file 1).

One of the main characteristics of the *Azospirillum* species is that they are able to fix nitrogen, i.e. convert N_2 to ammonia, by means of the nitrogenase enzyme complex. The main genes implicated in this process are known as *nif* genes, and they are highly conserved among nitrogen-fixing proteobacteria [30].

A preliminary BLAST analysis showed that the *A. amazonense nif* genes exhibit high similarity with genes of some species of the order Rhizobiales (Additional file 1). Since these observations were unexpected, the phylogenetic history of *nifH* was reconstructed utilizing sequences from species of the orders Rhodospirillales and Rhizobiales. The resulting *nifH* tree (Figure 4) was clearly incongruent to the 16S rDNA tree (Figure 1): in the *nifH* tree, *A. amazonense* and *A. irakense* grouped with *Bradyrhizobium* sp. BTAi1, *Xanthobacter diazotrophicus* and *Azorhizobium caulinodans*, instead of grouping with other *Azospirillum* species.

The genetic organization of the *nif* genes between the Rhodospirillales and Rhizobiales bacteria is somewhat homogeneous (Figure 5). As demonstrated in Figure 5, some features of the *nif* cluster of *A. amazonense* are exclusively similar to the homolog cluster of the species *Bradyrhizobium* sp. BTAi1, such as the presence of three conserved hypothetical genes that are indicated by the numbers 2, 5 and 6. On the other hand, *Azospirillum* sp. B510 and *A. brasilense* have the *aerC* gene between the *nifHDK* and *nifENX* operons and the *draG* and *draT* genes in the upstream region of the *nifH* gene, features not shared with the *nif* cluster of *A. amazonense*. The genes *draG* and *draT* code for a post-translational control system of the nitrogenase, which are not present in *A. amazonense* and *R. centenum* [9,31].

The domain composition of the deduced NifE and NifU proteins is also evidence that the nitrogenase complex of *A. amazonense* is more distantly related to *Azospirillum* spp. (with the exception of *A. irakense*) than to some Rhizobiales species (i.e. those from the genus *Bradyrhizobium, Xanthobacter* and *Azorhizobium*). The NifE protein of *A. amazonense* (like the *Bradyrhizobium, Xanthobacter* and *Azorhizobium* proteins NifE) has a bacterioferritin-associated ferredoxin [2Fe-2S] binding domain (BFD), not present in the NifE protein from *A. brasilense* Sp245 and *Azospirillum* sp. B510, and *R. palustris* (Rhizobiales) (Additional file 2). Furthermore, its NifU protein is smaller than those of *A. brasilense* Sp245, *Azospirillum* sp. B510 and *R. palustris* CGA009 because it does not contain the BFD and an N-terminal NifU domain, which are present in the NifU protein of the latter species (Additional file 2).

Therefore, taken together, these results indicate that complex events occurred in the evolution of the nitrogenase system in these bacteria, for instance, horizontal gene transfers and/or gene duplication followed by differential gene loss, culminating in the current distribution of the *nif* genes among the members of these taxonomic orders.

Nitrogen fixation is a very energy-demanding process and it is not surprising that the nitrogenase system is elaborately regulated. In all diazotrophic species of the Proteobacteria examined so far, the transcriptional activator NifA and the sigma N alternative RNA polymerase-associated factor are the master regulators of nitrogen fixation genes [32]. Sequence motifs similar to the consensus region of sigma N and NifA binding sites are present upstream of the *nifH* gene, the homolog of

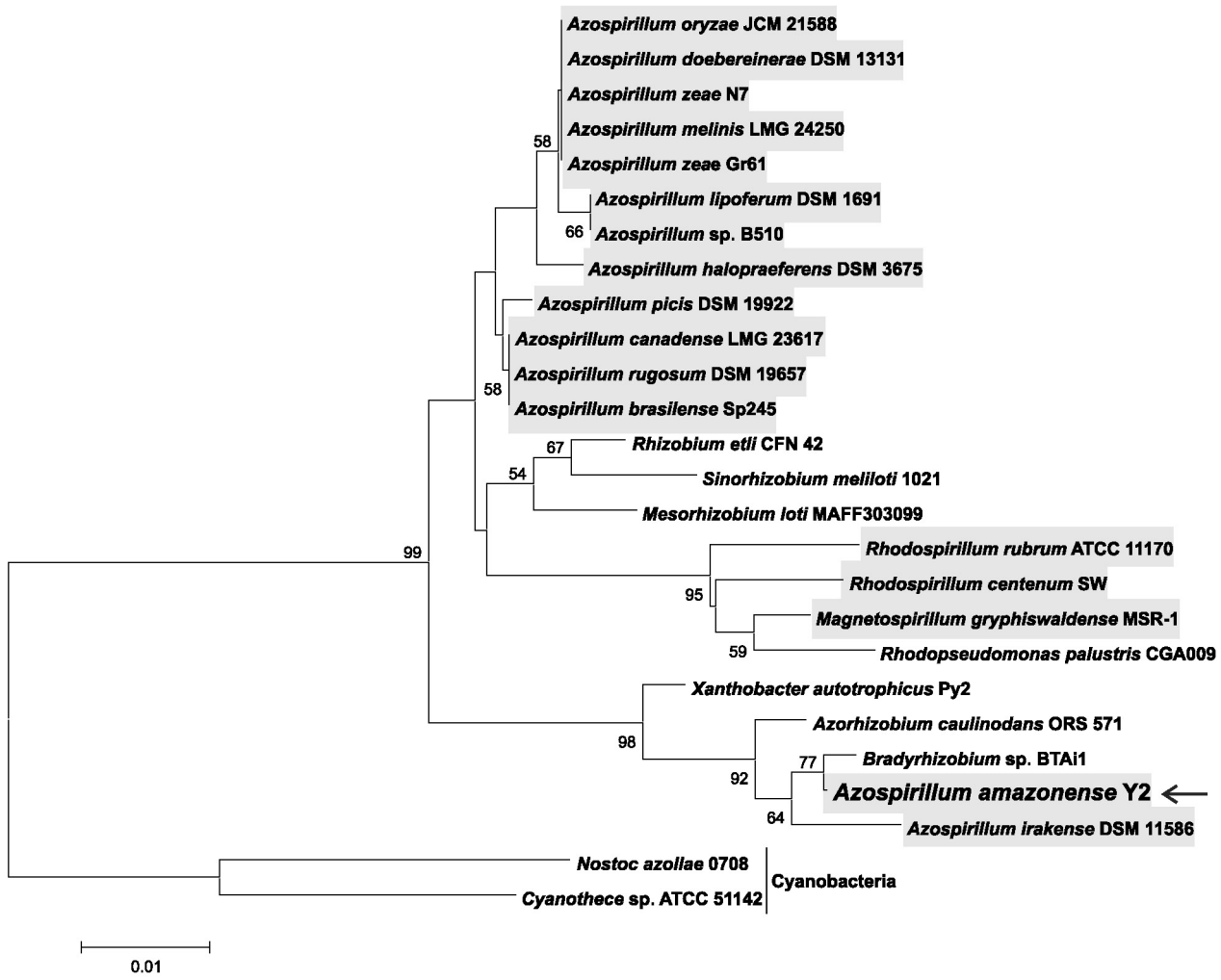


Figure 4 Phylogenetic tree based on the partial *nifH* gene from members of the orders Rhizobiales and Rhodospirillales. There were a total of 215 positions in the final dataset. Details are as shown in Figure 1.

"*orf2*" (indicated by the number 7 in Figure 5) of the *orf2nifUSVorf4* cluster from *A. brasilense* [33] and the *nifB* gene (Additional file 3).

Energy production and conversion for nitrogen fixation

Nitrogen fixation demands the systematic action of different genes. The *fix* genes are essential for this process, and they encompass different functional categories. These genes were found in the *A. amazonense* genome divided into three main clusters, namely *fixABCX, fixLJ-fixK* and *fixNOQP-fixGHIS*.

The *fixABCX* genes from *A. amazonense*, responsible for electron transfer to nitrogenase, are located downstream of the *nifW* gene (Figure 5). As seen in Figure 5, this genetic cluster is highly conserved among the analyzed species, and it is tightly associated with the *nif* cluster. Putative NifA and sigma N binding sites were found upstream of the operon *fixABCX* from *A. amazonense* (Additional file 3), indicating that these transcription factors could be key elements for the expression of this operon. This evidence is corroborated by reports showing that the operon *fixABCX* is regulated by the NifA protein in *A. brasilense* and *Rhizobium* spp. [34-36].

The *A. amazonense* genome also possesses a gene cluster including *fixLJ* and *fixK*. In symbiotic diazotrophs, the transcription of *fix* genes involves the oxygen-responsive FixLJ two-component system. The FixL protein, in the absence of oxygen, autophosphorylates and transfers the phosphate group to FixJ. Finally, the phosphorylated FixJ activates the expression of FixK, which activates the transcription of genes required for microaerobic growth [32].

Molecular nitrogen reduction requires high levels of energy under microaerobic conditions. The *fixNOQP* and *fixGHIS* genes encode membrane-bound

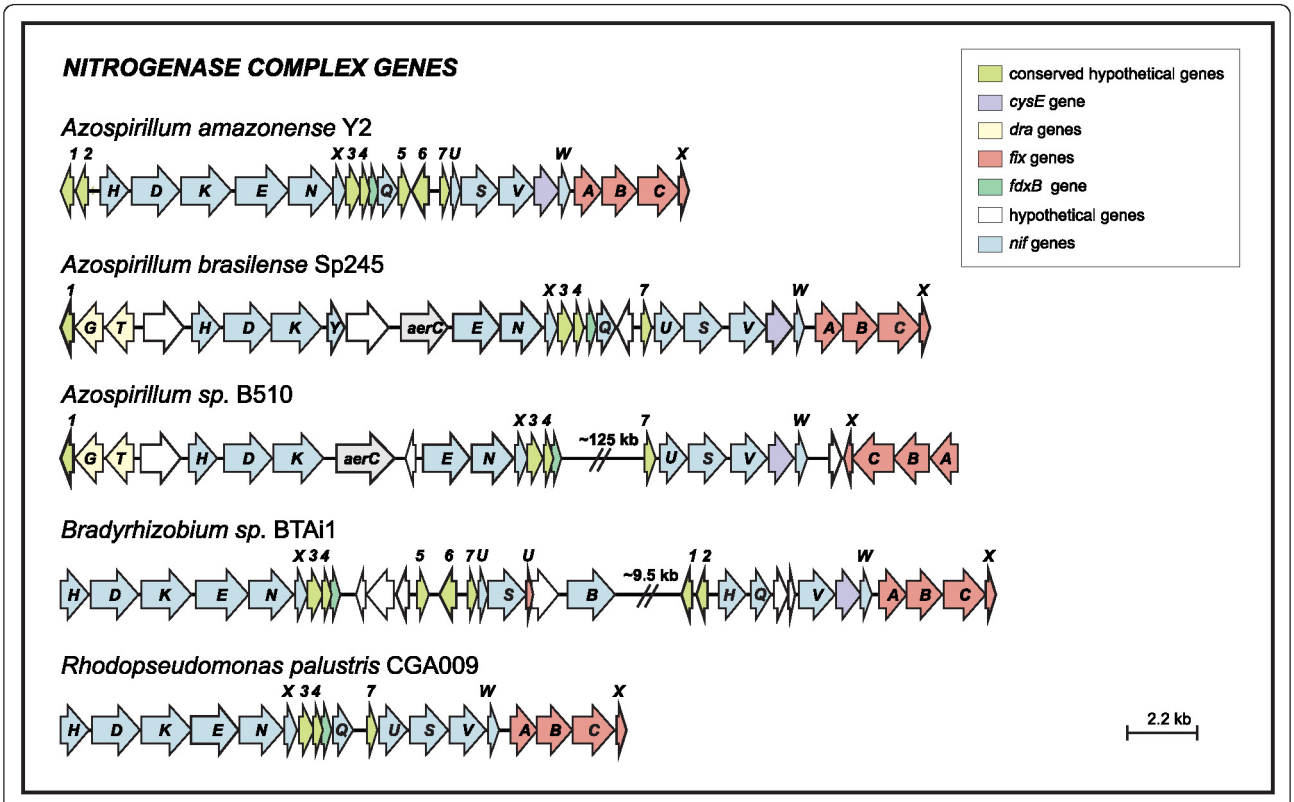


Figure 5 Organization of the nitrogen fixation gene cluster across different species. Arrows represent genes and their respective direction of transcription. Genes are colored as depicted in the upper box.

cytochrome c oxidase and the redox process-coupled cation pump, respectively, which are intimately implicated in respiration under microaerobic conditions, supplying energy for nitrogen fixation [37,38]. These clusters were found *in tandem* in the *A. amazonense* genome (Additional file 1), and they show identical organization in many diazotrophic α-Proteobacteria, like *R. centenum, A. brasilense* Sp245, *R. palustris* and *Bradyrhizobium* spp.

Nitrogen fixation forms molecular hydrogen (H_2) as a byproduct. Therefore, diazotrophic bacteria have several hydrogenase systems that are responsible for oxidizing molecular hydrogen to recover part of the energy expended during nitrogenase activity. Genes encoding for an uptake NiFe hydrogenase (*hupSL*) were identified in *A. amazonense* (Figure 6). Furthermore, the accessory proteins required for maturation of the subunits [39,40], encoded by the *hup* and *hyp* genes, are situated downstream of the genes encoding *hupSL (*Figure 6*)*. This organization resembles that found in members of the order Rhizobiales, where, in general, the *hup* and *hyp* genes are clustered (Figure 6), although their ordering is quite heterogeneous among the species. In contrast with this observation, comparisons with closely-related species showed that *Azospirillum* sp. B510 and *A. brasilense* Sp245 have their *hup* and *hyp* genes scattered across the genome. Moreover, the bacteria *R. centenum* does not have the hydrogenase gene cluster.

An integrated model relating the components discussed in this section with nitrogen metabolism is depicted in Figure 7, taken into consideration the similarity of *A. amazonense* genes to those from other well-known bacterial systems.

Phytohormone production

Species of the *Azospirillum* genus can positively influence plant growth and crop yield by the biosynthesis and secretion of indole-3-acetic acid (IAA) [2,41]. However, although *A. amazonense* strains are able to synthesize IAA [10], very little is known about the molecular mechanisms responsible for this process.

In *A. brasilense*, at least three pathways for IAA biosynthesis exist, two tryptophan-dependent pathways (indole-3-acetamide pathway (IAM) and indole-3-pyruvate pathway (IPyA)) and one tryptophan-independent pathway [3,42]. Similarly, the genome of *Azospirillum* sp. B510 contains genes responsible for the IAM pathway [7]. However, the *iaaM, iaaH* and *ipdC* genes, related to the IAM or IPyA pathways, were not located in the *A. amazonense* genome. Further analysis of the

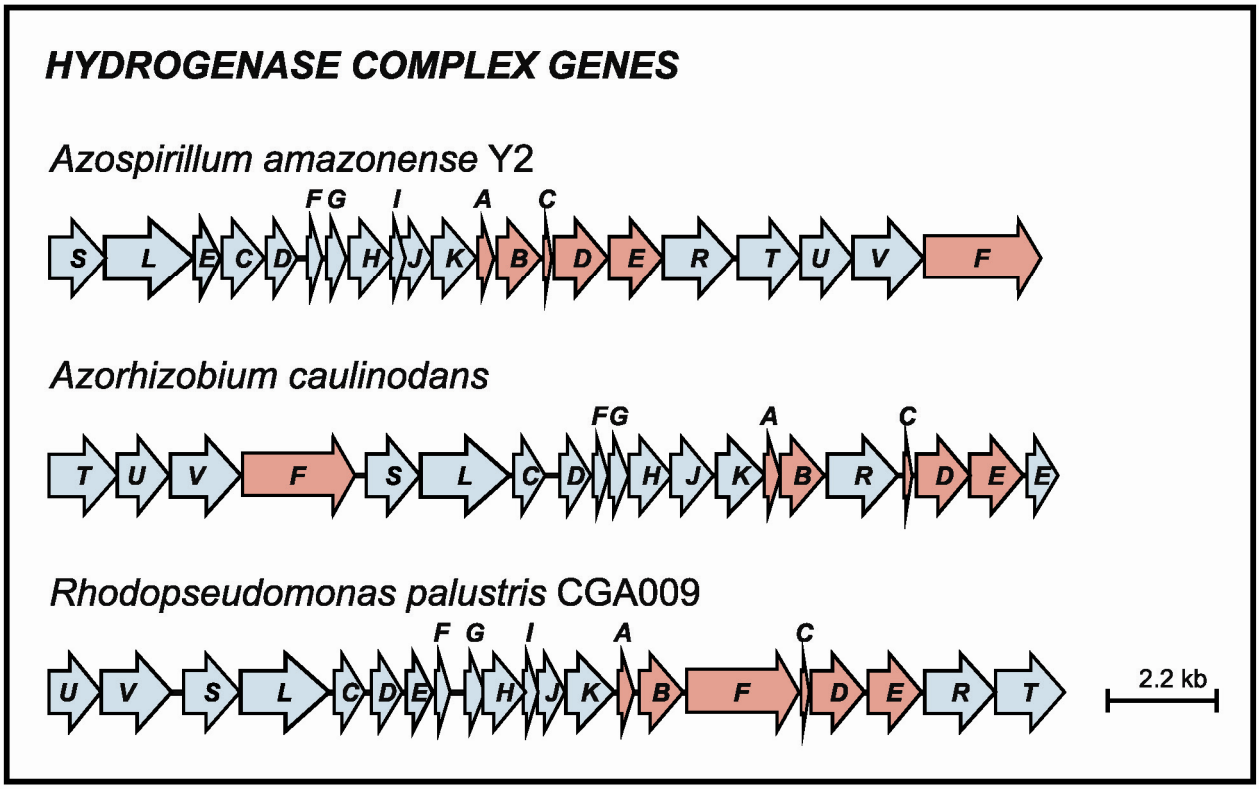


Figure 6 Organization of the hydrogenase gene cluster across different species from the orders Rhodospirillales and Rhizobiales. Arrows represent genes and their respective direction of transcription. *hup* genes are colored in red and *hyp* genes are colored in blue.

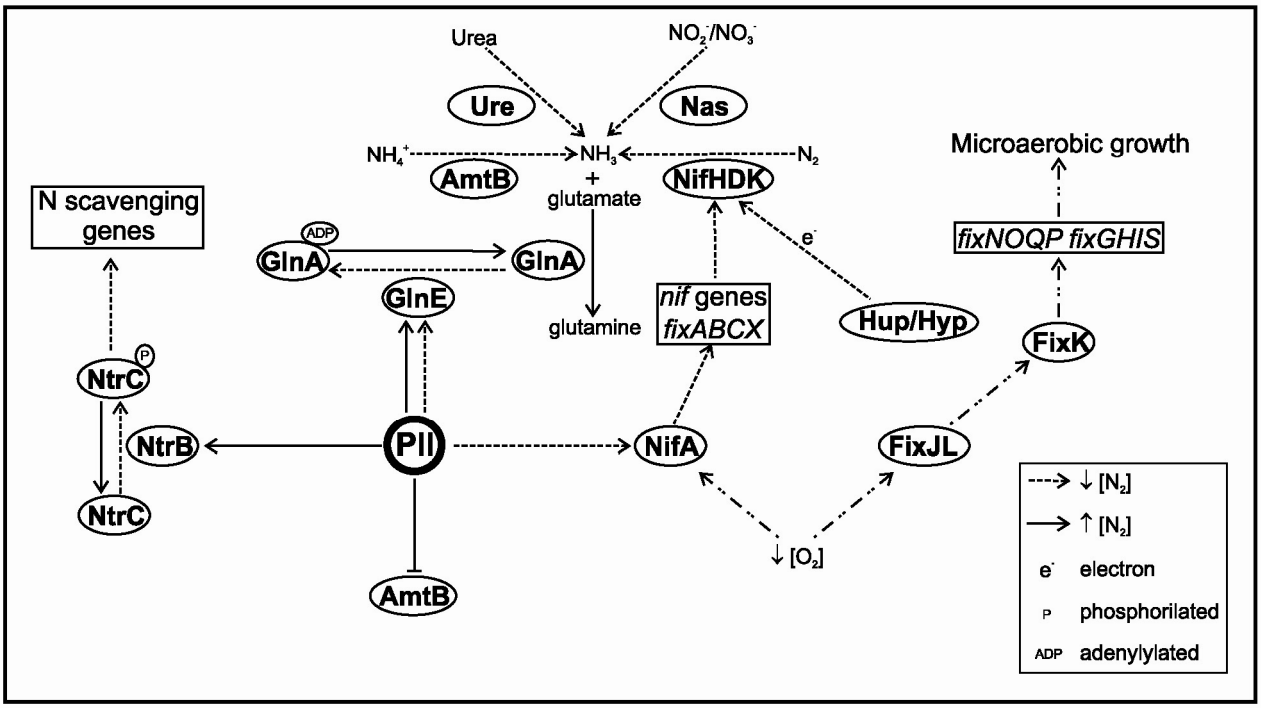


Figure 7 Integrated model correlating nitrogen metabolism and energy related-pathways for nitrogen fixation. The legend box indicates the correlation of each symbol with its respective meaning.

genome sequence of *A. amazonense* revealed a gene encoding a protein with about 70% similarity to nitrilases from plant species, like *Arabidopsis thaliana* and *Zea mays*, which catalyze the conversion of indole 3-acetonitrile to IAA [43,44]. Future studies may verify if this gene is implicated in IAA biosynthesis in *A. amazonense*.

Quorum sensing and biofilm formation

Quorum sensing is an intercellular signaling process implicated in the regulation of several traits of bacteria, notably antibiotic biosynthesis and biofilm formation. The archetype for quorum sensing regulation is the LuxIR system, which involves an acyl-homoserine lactone (AHL) synthase (LuxI homolog) and an AHL-dependent transcriptional regulator (LuxR homolog) [45]. The quorum sensing phenomenon of *Azospirillum* species is strain-specific and seems to regulate functions linked to rhizosphere competence and adaptation to plant roots [46].

The acyl-homoserine lactone (AHL) biosynthesis ability of forty *Azospirillum* strains (including *A. amazonense* Y2) was previously investigated, and only four *A. lipoferum* strains seemed of being capable of synthesizing these compounds [47]. However, the genome analysis of *A. amazonense* revealed the presence of genes encoding for LuxI and LuxR homologs proteins (Additional file 1). Therefore, these results indicate that *A. amazonense* Y2 could synthesize AHLs and respond to its presence in the environment. The genome of *A. amazonense* also presents a *Klebsiella pneumoniae ahlK* homolog [48], a predicted gene that codes for a putative homoserine lactonase (Additional file 1) implicated in AHL degradation. Since bacterial plant pathogens rely on quorum sensing mechanisms to infect plants [45], a study of *A. amazonense* homoserine lactonase activity on the deleterious activities of these pathogens would be relevant.

Extracellular polysaccharides are loosely bound to the cell surface and play an important role in bacterium-plant interactions through the firm and irreversible anchoring of cells to the plant roots [3]. In the *A. amazonense* genome, two genes, *noeJ* (mannose-6-phosphate isomerase, Additional file 1) and *noeL* (GDP-mannose 4,6-dehydratase, Additional file 1), which are related to extracellular polysaccharide biosynthesis and biofilm formation, were also found [49].

Chemotaxis/Motility

Different species of *Azospirillum* attach to and colonize plant root surfaces and these processes depend on chemotaxis. Azospirilla exhibit chemotaxis towards sugars, amino acids, organic acids and root exudates [50]. This ability offers the bacteria the advantage of moving towards favorable nutrient conditions.

Genes encoding for the central signal transduction pathway for chemotaxis (*che*) are present in nearly all motile bacteria. This signal transduction system is composed of the conserved *cheAWYBR* genes and a group of transmembrane chemoreceptors (known as MCPs or methyl-accepting proteins) that perceive environmental signals. Homologs of the *cheAWYBR* and MCP genes were identified in the *A. amazonense* genome (Additional file 1), and some *che* genes display similar organization to those found in the model organism *Escherichia coli* [51].

The *A. amazonense* Che1 gene cluster (*cheAWYBR*) revealed a conserved organization with the major chemotaxis gene cluster from *A. brasilense* Sp245, which modulates cell length and clumping behavior [52]. Apparently, this gene cluster also affects the production of exopolysaccharide and flocculation of *A. brasilense*. Other *che*-like genes (Che2 and Che3, Additional file 1) are present in the *A. amazonense* genome, probably encoding parallel signal transduction pathways that could have distinct functions, similar to those found in *R. centenum* [53,54].

It is also worth noting that genes encoding MCP domain proteins are spread throughout the *A. amazonense* genome. Some of them were classified accordingly to their similarities to *E. coli* chemoreceptors. All five types of *E. coli* MCP receptors were found in the *A. amazonense* genome (Additional file 1), and it is probable that some of them are related to cell motility by regulating the histidine kinase CheA that phosphorylates a response regulator, which in turn controls the rotational direction of the flagellar motor [55].

The flagellum is a key structure for the chemotactic response. In the *A. amazonense* genome, 39 flagellar genes were identified, and the majority of them are distributed among different gene clusters (Additional file 1). The reduced flagellar gene number in *A. amazonense* compared to those of the closely-related *A. brasilense* (79 annotated genes) and *R. centenum* (72 genes) species is in agreement with previous evidence indicating that *A. amazonense* synthesizes only the polar flagellum for swimming motility [56].

Homologs of the *A. brasilense* and *R. centenum* genes *fliFHN-motA-flbD-flhAF-fleN* [9,57] are present in *A. amazonense* (Additional file 1). The *A. amazonense* genome also contains other flagellar genes that display similar organization to those found in *Azospirillum* spp. and *R. centenum* (Figure 8).

Previous studies demonstrated that the transcription factor FlbD is related to the biosynthesis of lateral flagella in *A. brasilense* and the polar flagella in *R.*

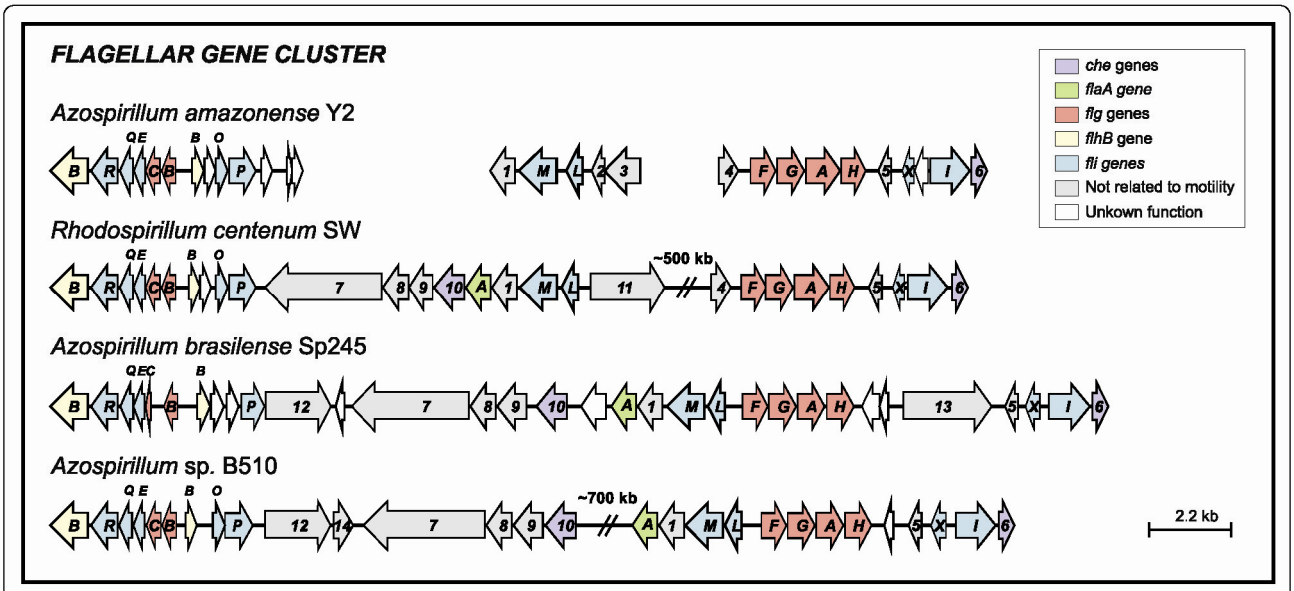


Figure 8 Organization of the flagellar gene cluster across different species. Arrows represent the genes and their respective direction of transcription. Genes are colored as depicted in the upper box. 1- ATPase involved in DNA repair, 2- glyoxalase/bleomycin resistance protein/dioxygenase, 3- twin-arginine translocation pathway signal, 4- peptidylprolyl isomerase, 5- RNA polymerase-binding protein DksA, 6- chemotactic signal-response protein CheL, 7- tetratricopeptide TPR_2 (cell wall/envelope/membrane biosynthesis), 8- chemotaxis MotB protein, 9- PPE-repeat protein, 10- chemotaxis protein CheZ, a-flaA, 11- transglutaminase-like domain protein, 12- K+ transporter, 13- O-methyltransferase, 14- beta-lg-H3/fasciclin.

centenum [9,57]. The presence of the *flbD* gene in the *A. amazonense* genome suggests that it could have a similar regulatory function as that found in *R. centenum*.

Transport, antibiotics resistance and lantibiotic production

Transport systems allow the uptake of nutrients and ions, excretion of end products of metabolism and communication between cells and the environment.

Several components of the cationic efflux pump, the ATP-Binding Cassette (ABC) transporter superfamily, the Major Facilitator Superfamily (MFS) and the TonB-dependent transporters were identified in the *A. amazonense* genome.

TonB-dependent transport is a mechanism of active uptake across the outer membrane normally related to iron uptake, signal transduction and environmental perception [58,59]. *A. amazonense* has a high number of genes encoding TonB receptors (Additional file 1) when compared with other diazotrophs such as *Azospirillum* sp. B510 (9 annotated genes) and *R. centenum SW* (32 annotated genes). Homology analyses suggest that a set of putative TonB receptors for specific substrates like ferrioxamine, cobalamin (B12 vitamin) and heme are present in this bacterium.

Among the different families of transporters, only two occur ubiquitously in all kingdoms of life: the Major Facilitator Superfamily (MFS) and the ATP-Binding Cassette (ABC) superfamily, representing the largest and most distributed families of transmembrane proteins. MFS proteins are single polypeptide secondary carriers that utilize uniport, symport or antiport mechanisms to transport various small substrates [60]. The ABC transporter proteins utilize energy from adenosine triphosphate (ATP) hydrolysis to carry out the uptake of essential nutrients and/or the extrusion of toxic substances [61]. In the *A. amazonense* genome, several genes encoding for putative MFS and ABC transporters were identified, and they could be implicated in the transport of a wide range of putative substrates (Additional file 1).

Multidrug resistance (MDR) transporters increase drug excretion through an efflux pump, which expels a wide variety of toxic products from the cell, playing a central role in bacterial drug resistance. The MDR transporters belong to various transporter families [62]. In *E. coli*, the transport of diverse substrates out of the cell by the AcrAB-TolC efflux transporter can confer broad resistance to antibiotics [63]. The *acrA* and *acrB* genes normally form an operon whose transcription is regulated by the *acrR* gene product, and are found in the *A. amazonense* genome.

Putative drug resistance transporters of the QacA subfamily were found in the *A. amazonense* genome (Additional file 1), which could confer resistance to monovalent and bivalent cationic lipophilic antiseptics

and disinfectants such as quaternary ammonium compounds [64].

Further analysis revealed that, in addition to the MDR transporters, a set of genes whose products could be related to specific antimicrobial resistance are present in the *A. amazonense* genome. Genes that code for penicillin, glyoxalase/bleomycin and tetracycline resistance are also present in the *A. amazonense* genome (Additional file 1). These findings corroborate the experimental data that shows that *A. amazonense* is tolerant to tetracycline and resistant to penicillin [6].

Lantibiotics are peptide-derived antibacterial substances produced by some bacteria, and are characterized by the presence of unusual amino acids like lanthionines and dehydrated amino acids [65]. Lantibiotic biosynthesis is frequently coregulated as part of a stress response when cells enter the late-log or stationary phase [66]. Most lantibiotics exert their antibiotic effect by either forming pores in the target cell membrane or by inhibiting cell wall synthesis, and many lantibiotics are bactericidal against a variety of Gram-positive bacteria [67]. Genes related to lantibiotic synthesis were found in *A. amazonense* genome (Additional file 1).

The genes implicated in antibiotic resistance and in lantibiotic production are probably essential for successful establishment of this microorganism in the soil due to constant contact with niche competitors like fungi and other bacteria.

Bacteriophytochrome

The *A. amazonense* genome also harbors a bacteriophytochrome gene (Additional file 1). Similarly, *Azospirillum* sp. B510 has two genes that code for distinct types of bacteriophytochromes [7], and *Bradyrhizobium* sp. possesses three bacteriophytochrome genes [68]. In plants, phytochromes regulate the metabolic response to the light environment, but a variety of functions is found in other organisms [69]. The bacteriophytochromes in *R. palustris* regulate the biosynthesis of the photosynthetic apparatus [70], while in *Deinococcus radiodurans* and *R. centenum*, they regulate pigment biosynthesis [71]. Subsequent studies must be carried out to understand the role of the bacteriophytochrome in *A. amazonense* physiology.

Conclusion

In order to thrive, bacteria must adapt readily to environmental shifts by means of a wide variety of genotypic and phenotypic accommodations [72]. The rhizosphere is a good example of a dynamic environment, where fluctuations in its biological and chemical activities demand an appropriate response from its inhabitants. The species *A. amazonense* is a free-living plant growth-promoting rhizobacterium that is found in association with plants of agricultural importance. In this study, we identified a series of *A. amazonense* genes that could be essential for adaptation to the competitive environment of the rhizosphere. Its wide genetic repertoire confers a versatile metabolism (e.g. the ability to use different carbon and nitrogen sources), as well as different mechanisms of perceiving and exploring its surroundings. These characteristics could directly influence plant growth, for instance, by providing nitrogen and stimulatory compounds to plants. Another important finding was the greater similarity of some genes, e.g. nitrogenase and RubisCO genes, to genes of members of the order Rhizobiales than to genes from other *Azospirillum* species. This evidence illustrates the genetic plasticity of this species and indicates that evolutionary phenomena like horizontal gene transfer could be fundamental for adaptation to its environment.

The major impact of this work will be to guide subsequent studies for a better understanding of the potential of *A. amazonense*.

Methods

Bacterial strain, culture conditions, and DNA isolation

A. amazonense Y2 (ATCC 35120) was cultured in M79 medium (10 g/L of sucrose as a carbon source, 0.1 g/L of K_2HPO_4, 0.4 g/L of KH_2PO_4, 0.2 g/L of $MgCl_2.7H_2O$, 0.1 g/L of NaCl, 0.4 g/L of yeast extract, pH 6.5) [73] with shaking at 150 rpm and 35°C for 18 hours. Genomic DNA was isolated as described by Wilson [74]. The quality of the isolated genomic DNA was assessed by agarose gel electrophoresis.

Genome sequencing, assembly, draft annotation

Total genomic DNA was sequenced using the Roche 454 pyrosequencing platform following the manufacturer's instructions. The contigs were assembled using Newbler software version 2.3 with the default parameters. The estimated coverage of the genome was 35x. Some gaps present in the genes of interest were filled in by sequencing PCR fragments obtained from genomic DNA.

The annotation and analysis of the sequences were carried out using the System for Automated Bacterial Integrated Annotation (SABIA) [75]. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AFBX00000000. The version described in this paper is the first version, AFBX01000000.

Phylogenetic analysis

Gene sequences were retrieved from GenBank. Most of the 16S rDNA sequences were retrieved from the Ribosomal Database Project [http://rdp.cme.msu.edu/] [76,77]. The accession numbers of the sequences utilized

in the phylogenetic reconstructions are listed in Additional file 4.

Multiple sequence alignments were performed using MUSCLE version 3.8 [78] and CLUSTALW (built into the MEGA 4 software) [79]. Phylogenetic trees were inferred using the neighbor-joining method (1000 bootstrap replicates) by the MEGA 4 software [80]. The evolutionary distances were computed using the Maximum Composite Likelihood method for the nucleotide sequences and the Jones-Taylor-Thornton (JTT) matrix-based method for the amino acid sequences. All positions containing gaps and missing data were eliminated from the datasets (complete deletion option).

Additional material

Additional file 1: Supplementary table 1. Genes of *Azospirillum amazonense* described throughout the study.

Additional file 2: Supplementary Figure 1. Domain composition of the NifE and NifU proteins among bacteria from the orders Rhodospirillales and Rhizobiales. The oblong boxes represent protein domains which are colored according to the description in the legend.

Additional file 3: Supplementary table 2. Putative sigma N and NifA binding sites occurring upstream of some *nif* and *fix* genes.

Additional file 4: Supplementary table 3. Accession numbers of the sequences utilized in the phylogenetic reconstructions.

Acknowledgements and Funding

We especially thank Dr. Augusto Schrank for the valuable effort that enabled the realization of this work. We also thank the Empresa Brasileira de Pesquisa Agropecuária - Centro Nacional de Pesquisa de Agrobiologia (EMBRAPA-CNPAB) for providing the *A. amazonense* Y2 strain. This work was supported by grants from the Brazilian National Research Council (CNPq) and the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS). FHS, RC, LAR and FMS received scholarships from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Author details

[1]Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Campus do Vale, Porto Alegre, RS, Brazil. [2]Laboratório Nacional de Computação Científica (LNCC), Petrópolis, RJ, Brazil. [3]Departamento de Biologia Molecular e Biotecnologia - Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, Campus do Vale, RS, Brazil.

Authors' contributions

ISS and ATRV conceived of and coordinated the study. FHS and RC extracted the genomic DNA. LGPA carried out the draft genome sequencing and assembly. FHS, LGPA, RC, LAR, FMS, MRSM and ISS performed the draft genome annotation. FHS, MRSM and RC carried out the comparative analyses. FHS carried out the phylogenetic analyses and created the illustrations. FHS, RC, LAR, FMS, MRSM and ISS analyzed the results. FMS and LAR wrote some sections of the manuscript. FHS and ISS wrote the manuscript. All authors read and approved the final manuscript.

Received: 29 April 2011 Accepted: 12 August 2011
Published: 12 August 2011

References

1. Bashan Y, De-Bashan LE: **How the plant growth-promoting bacterium *Azospirillum* promotes plant growth - a critical assessment.** *Adv Agron* 2010, **108**:77-136.
2. Bashan Y, Holguin G, De-Bashan LE: ***Azospirillum*-plant relationships: physiological, molecular, agricultural, and environmental advances (1997-2003).** *Can J Microbiol* 2004, **50**:521-577.
3. Steenhoudt O, Vanderleyden J: ***Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects.** *FEMS Microbiol Rev* 2000, **24**:487-506.
4. Lavrinenko K, Chernousova E, Gridneva E, Dubinina G, Akimov V, Kuever J, Lysenko A, Grabovich M: ***Azospirillum thiophilum* sp. nov., a novel diazotrophic bacterium isolated from a sulfide spring.** *Int J Syst Evol Microbiol* 2010, **60**:2832-2837.
5. Baldani JI, Baldani VLD: **History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience.** *An Acad Bras Ciênc* 2005, **77**:549-579.
6. Magalhães FMM, Baldani JI, Souto SM, Kuykendall JR, Döbereiner J: **A new acid-tolerant *Azospirillum* species.** *An Acad Bras Ciênc* 1983, **55**:417-430.
7. Kaneko T, Minamisawa K, Isawa T, Nakatsukasa H, Mitsui H, Kawaharada Y, Nakamura Y, Watanabe A, Kawashima K, Ono A, Shimizu Y, Takahashi C, Minami C, Fujishiro T, Kohara M, Katoh M, Nakazaki N, Nakayama S, Yamada M, Tabata S, Sato S: **Complete genomic structure of the cultivated rice endophyte *Azospirillum* sp. B510.** *DNA Res* 2010, **17**:37-50.
8. ***Azospirillum brasilense* Sp245.** [http://genome.ornl.gov/microbial/abra/19sep08/].
9. Lu YK, Marden J, Han M, Swingley WD, Mastrian SD, Chowdhury SR, Hao J, Helmy T, Kim S, Kurdoglu AA, Matthies HJ, Rollo D, Stothard P, Blankenship RE, Bauer CE, Touchman JW: **Metabolic flexibility revealed in the genome of the cyst-forming alpha-1 proteobacterium *Rhodospirillum centenum*.** *BMC genomics* 2010, **11**:325.
10. Rodrigues EP, Rodrigues LS, de Oliveira ALM, Baldani VLD, Teixeira KRS, Urquiaga S, Reis VM: ***Azospirillum amazonense* inoculation: effects on growth, yield and N_2 fixation of rice (*Oryza sativa* L.).** *Plant Soil* 2008, **302**:249-261.
11. Martin-Didonet CC, Chubatsu LS, Souza EM, Kleina M, Rego FG, Rigo LU, Yates MG, Pedrosa FO: **Genome structure of the genus *Azospirillum*.** *J Bacteriol* 2000, **182**:4113-4116.
12. Stoffels M, Castellanos T, Hartmann A: **Design and application of new 16S rRNA-targeted oligonucleotide probes for the *Azospirillum-Skermanella-Rhodocista*-cluster.** *Syst Appl Microbiol* 2001, **24**:83-97.
13. Fani R, Bandi C, Bazzicalupo M, Ceccherini MT, Fancelli S, Gallori E, Gerace L, Grifoni A, Miclaus N, Damiani G: **Phylogeny of the genus *Azospirillum* based on 16S rDNA sequence.** *FEMS Microbiol Lett* 1995, **129**:195-200.
14. Zhang D, Yang H, Zhang W, Huang Z, Liu SJ: ***Rhodocista pekingensis* sp. nov., a cyst-forming phototrophic bacterium from a municipal wastewater treatment plant.** *Int J Syst Evol Microbiol* 2003, **53**:1111-1114.
15. Kawasaki H, Hoshino Y, Kuraishi H, Yamasato K: ***Rhodocista centenaria* gen. nov., sp. nov., a cyst-forming anoxygenic photosynthetic bacterium and its phylogenetic position in the Proteobacteria alpha group.** *J Gen Appl Microbiol* 1992, **38**:541-551.
16. Martínez-Drets G, Fabiano E, Cardona A: **Carbohydrate catabolism in *Azospirillum amazonense*.** *Appl Environ Microbiol* 1985, **50**:183-185.
17. Faure D, Desair J, Keijers V, Bekri MA, Proost P, Henrissat B, Vanderleyden J: **Growth of *Azospirillum irakense* KBC1 on the aryl beta-glucoside salicin requires either *salA* or *salB*.** *J Bacteriol* 1999, **181**:3003-3009.
18. Kadouri D, Jurkevitch E, Okon Y: **Involvement of the reserve material poly-beta-hydroxybutyrate in *Azospirillum brasilense* stress endurance and root colonization.** *Appl Environ Microbiol* 2003, **69**:3244-3250.
19. Badger MR, Bek EJ: **Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO2 acquisition by the CBB cycle.** *J Exp Bot* 2008, **59**:1525-1541.
20. Hartmann A, Baldani JI: **The genus *Azospirillum*.** In *The Prokaryotes*. Edited by: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E. Springer New York; 2006:115-140.
21. Hartmann A, Zimmer W: **Physiology of *Azospirillum*.** In *Azospirillum/Plant Associations*. Edited by: Okon Y. Boca Raton: CRC Press; 1994:15-39.
22. Leigh JA, Dodsworth JA: **Nitrogen regulation in bacteria and archaea.** *Annu Rev Microbiol* 2007, **61**:349-377.
23. Arcondéguy T, Jack R, Merrick M: **PII signal transduction proteins, pivotal players in microbial nitrogen control.** *Microbiol Mol Biol Rev* 2001, **65**:80-105.
24. Sant'Anna FH, Andrade DS, Trentini DB, Weber SS, Schrank IS: **Tools for genetic manipulation of the plant growth-promoting bacterium *Azospirillum amazonense*.** *BMC Microbiol* 2011, **11**:107.

25. Sant'Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC, Schrank IS: **The PII superfamily revised: a novel group and evolutionary insights.** *J Mol Evol* 2009, **68**:322-336.
26. Javelle A, Merrick M: **Complex formation between AmtB and GlnK: an ancestral role in prokaryotic nitrogen control.** *Biochem Soc Trans* 2005, **33**:170-172.
27. Van Dommelen A, Spaepen S, Vanderleyden J: **Identification of the glutamine synthetase adenylyltransferase of *Azospirillum brasilense*.** *Res Microbiol* 2009, **160**:205-212.
28. Atkinson MR, Ninfa AJ: **Characterization of the GlnK protein of *Escherichia coli*.** *Mol Microbiol* 1999, **32**:301-313.
29. Shingler V: **Signal sensory systems that impact σ(54) -dependent transcription.** *FEMS Microbiol Rev* 2011, **35**:425-440.
30. Raymond J, Siefert JL, Staples CR, Blankenship RE: **The natural history of nitrogen fixation.** *Mol Biol Evol* 2004, **21**:541-554.
31. Fu HA, Hartmann A, Lowery RG, Fitzmaurice WP, Roberts GP, Burris RH: **Posttranslational regulatory system for nitrogenase activity in *Azospirillum* spp.** *J Bacteriol* 1989, **171**:4679-4685.
32. Dixon R, Kahn D: **Genetic regulation of biological nitrogen fixation.** *Nat Rev Microbiol* 2004, **2**:621-631.
33. Frazzon J, Schrank IS: **Sequencing and complementation analysis of the *nifUSV* genes from *Azospirillum brasilense*.** *FEMS Microbiol Lett* 1998, **159**:151-158.
34. David M, Daveran ML, Batut J, Dedieu A, Domergue O, Ghai J, Hertig C, Boistard P, Kahn D: **Cascade regulation of *nif* gene expression in *Rhizobium meliloti*.** *Cell* 1988, **54**:671-683.
35. Miller SH, Elliot RM, Sullivan JT, Ronson CW: **Host-specific regulation of symbiotic nitrogen fixation in *Rhizobium leguminosarum* biovar *trifolii*.** *Microbiology* 2007, **153**:3184-3195.
36. Sperotto RA, Gross J, Vedoy C, Passaglia LM, Schrank IS: **The electron transfer flavoprotein *fixABCX* gene products from *Azospirillum brasilense* show a NifA-dependent promoter regulation.** *Curr Microbiol* 2004, **49**:267-273.
37. Preisig O, Anthamatten D, Hennecke H: **Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis.** *Proc Natl Acad Sci USA* 1993, **90**:3309-3313.
38. Preisig O, Zufferey R, Hennecke H: **The *Bradyrhizobium japonicum fixGHIS* genes are required for the formation of the high-affinity cbb3-type cytochrome oxidase.** *Arch Microbiol* 1996, **165**:297-305.
39. Casalot L, Rousset M: **Maturation of the [NiFe] hydrogenases.** *Trends Microbiol* 2001, **9**:228-237.
40. Manyani H, Rey L, Palacios JM, Imperial J, Ruiz-Argüeso T: **Gene products of the *hupGHIJ* operon are involved in maturation of the iron-sulfur subunit of the [NiFe] hydrogenase from *Rhizobium leguminosarum* bv. *viciae*.** *J Bacteriol* 2005, **187**:7018-7026.
41. Malhotra M, Srivastava S: **An *ipdC* gene knock-out of *Azospirillum brasilense* strain SM and its implications on indole-3-acetic acid biosynthesis and plant growth promotion.** *Antonie van Leeuwenhoek* 2008, **93**:425-433.
42. Spaepen S, Vanderleyden J, Remans R: **Indole-3-acetic acid in microbial and microorganism-plant signaling.** *FEMS Microbiol Rev* 2007, **31**:425-448.
43. Kriechbaumer V, Park WJ, Piotrowski M, Meeley RB, Gierl A, Glawischnig E: **Maize nitrilases have a dual role in auxin homeostasis and beta-cyanoalanine hydrolysis.** *J Exp Bot* 2007, **58**:4225-4233.
44. Vorwerk S, Biernacki S, Hillebrand H, Janzik I, Muller A, Weiler EW, Piotrowski M: **Enzymatic characterization of the recombinant *Arabidopsis thaliana* nitrilase subfamily encoded by the NIT2/NIT1/NIT3-gene cluster.** *Planta* 2001, **212**:508-516.
45. Loh J, Pierson EA, Pierson L3, Stacey G, Chatterjee A: **Quorum sensing in plant-associated bacteria.** *Curr Opin Plant Biol* 2002, **5**:285-290.
46. Boyer M, Bally R, Perrotto S, Chaintreuil C, Wisniewski-Dyé F: **A quorum-quenching approach to identify quorum-sensing-regulated functions in *Azospirillum lipoferum*.** *Res Microbiol* 2008, **159**:699-708.
47. Vial L, Cuny C, Gluchoff-Fiasson K, Comte G, Oger PM, Faure D, Dessaux Y, Bally R, Wisniewski-Dyé F: **N-acyl-homoserine lactone-mediated quorum-sensing in *Azospirillum*: an exception rather than a rule.** *FEMS Microbiol Ecol* 2006, **58**:155-168.
48. Park SY, Lee SJ, Oh TK, Oh JW, Koo BT, Yum DY, Lee JK: **AhlD, an N-acylhomoserine lactonase in *Arthrobacter* sp., and predicted homologues in other bacteria.** *Microbiology* 2003, **149**:1541-1550.
49. Lerner A, Castro-Sowinski S, Valverde A, Lerner H, Dror R, Okon Y, Burdman S: **The *Azospirillum brasilense* Sp7 *noeJ* and *noeL* genes are involved in extracellular polysaccharide biosynthesis.** *Microbiology* 2009, **155**:4058-4068.
50. Zhulin IB, Armitage JP: **Motility, chemokinesis, and methylation-independent chemotaxis in *Azospirillum brasilense*.** *J Bacteriol* 1993, **175**:952-958.
51. Wadhams GH, Armitage JP: **Making sense of it all: bacterial chemotaxis.** *Nat Rev Mol Cell Biol* 2004, **5**:1024-1037.
52. Bible AN, Stephens BB, Ortega DR, Xie Z, Alexandre G: **Function of a chemotaxis-like signal transduction pathway in modulating motility, cell clumping, and cell length in the alphaproteobacterium *Azospirillum brasilense*.** *J Bacteriol* 2008, **190**:6365-6375.
53. Berleman JE, Bauer CE: **Involvement of a Che-like signal transduction cascade in regulating cyst cell development in *Rhodospirillum centenum*.** *Mol Microbiol* 2005, **56**:1457-1466.
54. Jiang ZY, Gest H, Bauer CE: **Chemosensory and photosensory perception in purple photosynthetic bacteria utilize common signal transduction components.** *J Bacteriol* 1997, **179**:5720-5727.
55. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA: **The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes.** *Annu Rev Cell Dev Biol* 1997, **13**:457-512.
56. Moens S, Michiels K, Keijers V, Van Leuven F, Vanderleyden J: **Cloning, sequencing, and phenotypic analysis of *laf1*, encoding the flagellin of the lateral flagella of *Azospirillum brasilense* Sp7.** *J Bacteriol* 1995, **177**:5419-5426.
57. Chang Y, Tang T, Li JL: **Isolation of a flagellar operon in *Azospirillum brasilense* and functional analysis of FlbD.** *Res Microbiol* 2007, **158**:521-528.
58. Koebnik R: **TonB-dependent trans-envelope signalling: the exception or the rule?** *Trends Microbiol* 2005, **13**:343-347.
59. Noinaj N, Guillier M, Barnard TJ, Buchanan SK: **TonB-dependent transporters: regulation, structure, and function.** *Annu Rev Microbiol* 2010, **64**:43-60.
60. Pao SS, Paulsen IT, Saier MH: **Major facilitator superfamily.** *Microbiol Mol Biol Rev* 1998, **62**:1-34.
61. Jones PM, George AM: **The ABC transporter structure and mechanism: perspectives on recent research.** *Cell Mol Life Sci* 2004, **61**:682-699.
62. Saier MH, Paulsen IT: **Phylogeny of multidrug transporters.** *Semin Cell Dev Biol* 2001, **12**:205-213.
63. Li XZ, Nikaido H: **Efflux-mediated drug resistance in bacteria: an update.** *Drugs* 2009, **69**:1555-1623.
64. Brown MH, Skurray RA: **Staphylococcal multidrug efflux protein QacA.** *J Mol Microbiol Biotechnol* 2001, **3**:163-170.
65. Chatterjee C, Paul M, Xie L, van der Donk WA: **Biosynthesis and mode of action of lantibiotics.** *Chem Rev* 2005, **105**:633-684.
66. Okuda Ki, Yanagihara S, Sugayama T, Zendo T, Nakayama J, Sonomoto K: **Functional significance of the E-Loop, a novel motif conserved in the lantibiotic-immunity ABC transport systems.** *J Bacteriol* 2010, **192**:2801-2808.
67. Cotter P, Hill C, Ross P: **Bacteriocins: developing innate immunity for food.** *Nat Rev Microbiol* 2005, **3**:777-788.
68. Jaubert M, Vuillet L, Hannibal L, Adriano JM, Fardoux J, Bouyer P, Bonaldi K, Fleischman D, Giraud E, Vermeglio A: **Control of peripheral light-harvesting complex synthesis by a bacteriophytochrome in the aerobic photosynthetic bacterium *Bradyrhizobium* strain BTAi1.** *J Bacteriol* 2008, **190**:5824-5831.
69. Rockwell NC, Su YS, Lagarias JC: **Phytochrome structure and signaling mechanisms.** *Annu Rev Plant Biol* 2006, **57**:837-858.
70. Giraud E, Zappa S, Vuillet L, Adriano JM, Hannibal L, Fardoux J, Berthomieu C, Bouyer P, Pignol D, Verméglio A: **A new type of bacteriophytochrome acts *in tandem* with a classical bacteriophytochrome to control the antennae synthesis in *Rhodopseudomonas palustris*.** *J Biol Chem* 2005, **280**:32389-32397.
71. Davis SJ, Vener AV, Vierstra RD: **Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic Eubacteria.** *Science* 1999, **286**:2517-2520.
72. Roszak DB, Colwell RR: **Survival strategies of bacteria in the natural environment.** *Microbiol Rev* 1987, **51**:365-379.

73. Fred EB, Waskman SA: *Laboratory Manual of General Microbiology with Special Reference to the Microorganisms of the Soil* New York: McGraw-Hill Book Company, Inc; 1928.
74. Wilson K: **Preparation of genomic DNA from bacteria.** In *Current Protocols in Molecular Biology*.. 1 edition. Edited by: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. New York: Wiley; 1997:-2.
75. Almeida LGP, Paixão R, Souza RC, Costa GC, arrientos FJA, antos MT, Imeida DF, asconcelos ATR: **A System for Automated Bacterial (genome) Integrated Annotation–SABIA.** *Bioinformatics* 2004, **20**:2832-2833.
76. Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM: **The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data.** *Nucleic Acids Res* 2007, **35**:D169-D172.
77. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM: **The Ribosomal Database Project: improved alignments and new tools for rRNA analysis.** *Nucleic Acids Res* 2009, **37**:D141-D145.
78. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic Acids Res* 2004, **32**:1792-1797.
79. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
80. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: molecular evolutionary fenetics analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**:1596-1599.

doi:10.1186/1471-2164-12-409
Cite this article as: Sant'Anna *et al.*: Genomic insights into the versatility of the plant growth-promoting bacterium *Azospirillum amazonense*. *BMC Genomics* 2011 **12**:409.

Additional file 1. Supplementary table 1. Genes of *A. amazonense* described throughout the study

ORFS

Id	Name	Product	Category	Classification
AZA02775		ABC-type transporter, ATPase and permease components	ABC transporter	
AZA06024		ABC transporter, ATP-binding protein	ABC transporter	
AZA07069		ABC transporter, ATP-binding protein	ABC transporter	
AZA09826		ABC transporter, ATP-binding protein	ABC transporter	
AZA15101		aliphatic sulfonate ABC transporter ATP-binding protein	ABC transporter	
AZA19151		ABC-type branched-chain amino acid transport ATPase component	ABC transporter	
AZA21566		ABC-type siderophore export system	ABC transporter	
AZA21803		ABC transporter related protein	ABC transporter	
AZA27760		ABC-type sugar transport system, periplasmic component	ABC transporter	
AZA36299		ABC-type export system, outer membrane channel protein	ABC transporter	
AZA36319	macB	macrolide transport protein ABC-type efflux carrier	ABC transporter	
AZA36612		ABC transporter, ATP-binding protein	ABC transporter	
AZA40587	mkl	ABC transporter, ribonucleotide transport ATP-binding protein	ABC transporter	
AZA44644		ABC transporter, permease protein	ABC transporter	
AZA53052	utrC	urea ABC transporter, permease protein	ABC transporter	
AZA55128		ABC transporters with duplicated ATPase domains	ABC transporter	
AZA62081		ABC transporter related protein	ABC transporter	
AZA62938		ABC transporter, ATP-binding protein, putative	ABC transporter	
AZA66683		ABC transporter permease	ABC transporter	
AZA70834		putative ABC transporter ATP-binding protein	ABC transporter	
AZA71708		peptide ABC transporter, periplasmic peptide-binding protein	ABC transporter	
AZA76751		ABC transporter, periplasmic substrate-binding protein	ABC transporter	
AZA83258		ABC-type amino acid transport/signal transduction systems	ABC transporter	
AZA83819		ABC transporter related protein	ABC transporter	
AZA87640		putative ABC transporter, ATP-binding protein	ABC transporter	
AZA88337		ABC transporter ATP-binding protein	ABC transporter	
AZA88760		ABC transporter, permease protein	ABC transporter	
AZA88809		ABC-type transporter, periplasmic binding protein	ABC transporter	
AZA88925		ABC transporter related protein	ABC transporter	
AZA89042		ABC transporter, permease protein	ABC transporter	
AZA89100		ABC transporter, permease protein	ABC transporter	
AZA89130		glutamine transport system permease protein (ABC family)	ABC transporter	

AZA89131		amino acid ABC transporter	ABC transporter
AZA89360		sugar ABC transporter, permease protein	ABC transporter
AZA89361		sugar ABC transporter, ATP-binding protein	ABC transporter
AZA89586		sugar ABC transporter, ATP-binding protein	ABC transporter
AZA89588	ycjP	ABC transporter permease protein ycjP	ABC transporter
AZA89589		sugar ABC transporter, substrate-binding protein	ABC transporter
AZA89605		ABC transporter	ABC transporter
AZA89688		ABC transporter ATP-binding protein	ABC transporter
AZA89696		macrolide ABC transporter ATP-binding/membrane protein	ABC transporter
AZA89734		ABC transporter	ABC transporter
AZA89773		n peptide/opine/nickel uptake ABC transporter ATP-binding protein	ABC transporter
AZA89994		ABC transporter, ATP-binding protein	ABC transporter
AZA89996		ABC transporter, permease protein	ABC transporter
AZA90133		ABC transporter permease protein, putative	ABC transporter
AZA90202		ABC-type nitrate/sulfonate/bicarbonate transport system	ABC transporter
AZA90215		ABC superfamily ATP binding cassette transporter	ABC transporter
AZA90416	utrA	urea ABC transporter, urea binding protein	ABC transporter
AZA90417	utrB	urea ABC transporter, permease protein UrtB	ABC transporter
AZA90418	utrD	urea ABC transporter, ATP-binding protein UrtD	ABC transporter
AZA90419		urea ABC transporter, ATP-binding protein UrtE	ABC transporter
AZA06086		acriflavin resistance protein	Antibiotic resistance
AZA06133		acriflavin resistance protein	Antibiotic resistance
AZA06350		Beta-lactamase	Antibiotic resistance
AZA07395		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA09088		multidrug resistance protein	Antibiotic resistance
AZA10909		acriflavin resistance protein	Antibiotic resistance
AZA11424		drug resistance transporter, EmrB/QacA subfamily	Antibiotic resistance
AZA16488		beta-lactamase class C	Antibiotic resistance
AZA20396	acyI	penicillin amidase	Antibiotic resistance
AZA25605		metal-dependent hydrolases of the beta-lactamase superfamily I	Antibiotic resistance
AZA34622		metallo-beta-lactamase	Antibiotic resistance
AZA35214		cation/multidrug efflux pump	Antibiotic resistance
AZA35257	acrB	acriflavin resistance protein	Antibiotic resistance
AZA37804		acriflavin resistance protein	Antibiotic resistance
AZA38070		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA39395	mrcA	penicillin-binding protein 1A	Antibiotic resistance
AZA39732		beta-lactamase, putative	Antibiotic resistance

AZA40019		putative metallo-beta-lactamase	Antibiotic resistance
AZA44895		acriflavin resistance protein	Antibiotic resistance
AZA46788	mrdA	penicillin-binding protein 2	Antibiotic resistance
AZA47678		fusaric acid resistance protein	Antibiotic resistance
AZA48055	yegN	probable multidrug efflux pump, AcrB/AcrD/AcrF family	Antibiotic resistance
AZA56169		beta-lactamase	Antibiotic resistance
AZA59089		acriflavin resistance protein	Antibiotic resistance
AZA59335		acriflavin resistance protein	Antibiotic resistance
AZA62861		penicillin-binding protein, 1A family	Antibiotic resistance
AZA63807	ampC	putative beta-lactamase precursor (cephalosporinase)	Antibiotic resistance
AZA66998		beta-lactamase	Antibiotic resistance
AZA67294		drug resistance transporter, EmrB/QacA subfamily	Antibiotic resistance
AZA68914		beta-lactamase	Antibiotic resistance
AZA72319		probable tetracycline 6-hydroxylase	Antibiotic resistance
AZA74640		drug resistance transporter, EmrB/QacA subfamily	Antibiotic resistance
AZA74663		multidrug resistance protein	Antibiotic resistance
AZA75319		beta-lactamase	Antibiotic resistance
AZA75353		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA79411		outer membrane multidrug efflux protein	Antibiotic resistance
AZA84109	acrB	acriflavin resistance protein	Antibiotic resistance
AZA85715		multidrug resistance efflux system protein	Antibiotic resistance
AZA87534		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA87557	proP	major facilitator superfamily (MFS) multidrug efflux pump	Antibiotic resistance
AZA87630		putative glyoxylase/belomycin resistance protein	Antibiotic resistance
AZA87694	mdtN	multidrug resistance protein MdtN	Antibiotic resistance
AZA87717		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA87935		beta-lactamase	Antibiotic resistance
AZA87971		glyoxalase/bleomycin resistance protein/dihydroxybiphenyl	Antibiotic resistance
AZA88541	acrA	efflux transporter	Antibiotic resistance
AZA88558	mdtA	multidrug resistance protein mdtA	Antibiotic resistance
AZA88732	emrB	multidrug resistance protein B	Antibiotic resistance
AZA88734	emrA	multidrug resistance protein A	Antibiotic resistance
AZA89502		metallo-beta-lactamase family protein	Antibiotic resistance
AZA89575		daunorubicin/doxorubicin resistance ATP-binding protein drrA	Antibiotic resistance
AZA89640		multidrug resistance protein	Antibiotic resistance
AZA89643		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA89667		non-motile and phage-resistance protein	Antibiotic resistance

AZA89788		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA89961	drrA	daunorubicin resistance ATP-binding protein	Antibiotic resistance
AZA90138	bcr	Bcr/CflA subfamily drug resistance transporter	Antibiotic resistance
AZA90166		glyoxalase/bleomycin resistance protein/dioxygenase	Antibiotic resistance
AZA90169		beta-lactamase	Antibiotic resistance
AZA90232		multidrug resistance efflux pump	Antibiotic resistance
AZA90322		multidrug efflux pump	Antibiotic resistance
AZA90365		tetracycline-efflux transporter	Antibiotic resistance
AZA90492		beta-lactamase	Antibiotic resistance
AZA33581	bphB	bacteriophytochrome protein	Bacteriophytochrome
AZA38997	noeJ	mannose-1-phosphate guanylyltransferase	Biofilm formation
AZA88211	noeL	GDP-mannose 4,6-dehydratase	Biofilm formation
AZA80303	cbbE	ribulose-phosphate 3-epimerase	Carbon Fixation
AZA80325	cbbL	ribulose-bisphosphate carboxylase large chain	Carbon Fixation
AZA80388	cbbA	fructose-bisphosphate aldolase	Carbon Fixation
AZA80411	cbbT	transketolase (TK)	Carbon Fixation
AZA90487	cbbP	Phosphoribulokinase	Carbon Fixation
AZA90488	cbbR	transcriptional regulator, LysR family	Carbon Fixation
AZA90630	cbbF	fructose-1,6-bisphosphatase	Carbon Fixation
AZA90631	cbbS	ribulose bisphosphate carboxylase small chain	Carbon Fixation
AZA90632	cbbX	CbbX protein	Carbon Fixation
AZA89881	edd	6-phosphogluconate dehydratase	Entner-Doudoroff
AZA90627		KHG/KDPG aldolase	Entner-Doudoroff
AZA88036	fbaB	fructose-bisphosphate aldolase	Glycolysis
AZA90645		putative homoserine lactonase	Homoserine lactone degradation
AZA87454	hupC	Ni/Fe-hydrogenase 1 B-type cytochrome subunit	Hydrogen metabolism
AZA90297	hupD	hydrogenase expression/formation protein hupD	Hydrogen metabolism
AZA89067	hupE	urease accessory protein UreJ	Hydrogen metabolism
AZA89066	hupF	hydrogenase expression/formation protein hupF	Hydrogen metabolism
AZA89065	hupG	hydrogenase formation protein HupG	Hydrogen metabolism
AZA90479	hupH	hydrogenase formation protein	Hydrogen metabolism
AZA87443	hupI	rubredoxin	Hydrogen metabolism
AZA89062	hupJ	hydrogenase expression/formation protein	Hydrogen metabolism
AZA89061	hupK	hydrogenase expression/formation protein HupK	Hydrogen metabolism
AZA83590	hupL	hydrogenase protein large subunit	Hydrogen metabolism
AZA89053	hupR	hydrogenase transcriptional regulatory protein hupR	Hydrogen metabolism
AZA90299	hupS	hydrogenase small subunit	Hydrogen metabolism

AZA90643	hupT	HupT protein	Hydrogen metabolism
AZA90718	hupU	nickel-dependent hydrogenase small subunit	Hydrogen metabolism
AZA90720	hupV	nickel-dependent hydrogenase, large subunit	Hydrogen metabolism
AZA89059	hypA	hydrogenase nickel incorporation protein hypA	Hydrogen metabolism
AZA87413	hypB	hydrogenase accessory protein HypB	Hydrogen metabolism
AZA89058	hypC	hydrogenase expression/formation protein hypC	Hydrogen metabolism
AZA87446	hypD	hydrogenase expression/formation protein	Hydrogen metabolism
AZA89055	hypE	hydrogenase expression/formation protein HypE	Hydrogen metabolism
AZA90067	hypF	NiFe hydrogenase maturation protein HypF	Hydrogen metabolism
AZA12397		lanthionine synthetase C family protein	Lantibiotic biosynthesis
AZA81926	sunT	bacteriocin/lantibiotic ABC transporter	Lantibiotic biosynthesis
AZA85469	sunT	ABC-type bacteriocin/lantibiotic exporter	Lantibiotic biosynthesis
AZA06501		Major Facilitator Transporter	Major Facilitator Superfamily (MFS)
AZA07574		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA07803		MFS transporter	Major Facilitator Superfamily (MFS)
AZA15878		major facilitator superfamily protein	Major Facilitator Superfamily (MFS)
AZA16908		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA17681		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA29285		major facilitator family transporter	Major Facilitator Superfamily (MFS)
AZA57799		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA58482		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA59512		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA60508		major facilitator family transporter	Major Facilitator Superfamily (MFS)
AZA61555		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA75342		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA76479		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA79528		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA82255		MFS permease	Major Facilitator Superfamily (MFS)
AZA87336		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA88247		MFS transport system component	Major Facilitator Superfamily (MFS)
AZA88274		Major Facilitator Superfamily (MFS)	Major Facilitator Superfamily (MFS)
AZA88393		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA88880		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA88891		MFS transporter	Major Facilitator Superfamily (MFS)
AZA88905		major facilitator family transporter	Major Facilitator Superfamily (MFS)
AZA88994		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA89165	hsrA	MFS general substrate transporter	Major Facilitator Superfamily (MFS)

AZA89178		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA89223		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA89438	rmrB	major facilitator superfamily protein	Major Facilitator Superfamily (MFS)
AZA89613		putative major facilitator superfamily (MFS) transporter	Major Facilitator Superfamily (MFS)
AZA89751		major facilitator superfamily protein	Major Facilitator Superfamily (MFS)
AZA89765		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA89831		Major facilitator superfamily MFS_1 precursor	Major Facilitator Superfamily (MFS)
AZA89991		major facilitator family transporter	Major Facilitator Superfamily (MFS)
AZA90068		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA90091		major facilitator family transporter	Major Facilitator Superfamily (MFS)
AZA90263		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA90356		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA90509		major facilitator superfamily MFS_1	Major Facilitator Superfamily (MFS)
AZA90594		major facilitator transporter	Major Facilitator Superfamily (MFS)
AZA55000	fixL	PAS/PAC sensor signal transduction histidine kinase	Microaerophilic respiration
AZA87394	fixK	CRP/FNR family transcriptional regulator	Microaerophilic respiration
AZA87406	fixH	cation pump-linked membrane protein fixH	Microaerophilic respiration
AZA88457	fixU	FixU protein	Microaerophilic respiration
AZA89846	fixN	cytochrome c oxidase subunit 1,cbb3-type	Microaerophilic respiration
AZA89847	fixO	cytochrome c oxidase, cbb3-type, subunit II	Microaerophilic respiration
AZA89849	fixP	cytochrome c oxidase, Cbb3-type, subunit III	Microaerophilic respiration
AZA89851	fixG	cytochrome c oxidase cbb3 type accessory protein	Microaerophilic respiration
AZA90642	fixQ	FixQ protein	Microaerophilic respiration
AZA90698	fixJ	response regulator receiver protein	Microaerophilic respiration
AZA90710	fixS	cytochrome oxidase maturation protein cbb3-type	Microaerophilic respiration
AZA90712	fixJ	transcriptional regulatory protein	Microaerophilic respiration
AZA90655	nasE	nitrite reductase small subunit	Nitrate/nitrite metabolism
AZA90025	cysE	serine O-acetyltransferase	Nitrogen fixation
AZA90635	fdxA	ferredoxin	Nitrogen fixation
AZA90029	fixA	electron transfer flavoprotein alpha/beta-subunit	Nitrogen fixation
AZA90628	fixB	FixB protein	Nitrogen fixation
AZA82897	fixC	electron-transferring-flavoprotein dehydrogenase	Nitrogen fixation
AZA90637	fixX	putative ferredoxin protein, FixX	Nitrogen fixation
AZA39368	nifA	transcriptional regulator, NifA	Nitrogen fixation
AZA65617	nifB	nitrogenase cofactor biosynthesis protein	Nitrogen fixation
AZA00090	nifD	nitrogenase molybdenum-iron protein alpha chain	Nitrogen fixation
AZA87419	nifE	nitrogenase MoFe cofactor biosynthesis	Nitrogen fixation

AZA90011	nifH	nitrogenase reductase	Nitrogen fixation
AZA00102	nifK	nitrogenase molybdenum-iron protein beta chain	Nitrogen fixation
AZA00120	nifN	nitrogenase molybdenum-iron cofactor biosynthesis protein	Nitrogen fixation
AZA90400	nifQ	nitrogen fixation protein NifQ	Nitrogen fixation
AZA90300	nifS	cysteine desulfurase	Nitrogen fixation
AZA90636	nifU	NifU protein	Nitrogen fixation
AZA90024	nifV	homocitrate synthase	Nitrogen fixation
AZA90027	nifW	nitrogenase-stabilizing/protective protein nifW	Nitrogen fixation
AZA90014	nifX	nitrogenase molybdenum-iron protein	Nitrogen fixation
AZA00208		conserved hypothetical protein (4)	Nitrogen fixation
AZA90016		conserved hypothetical protein (3)	Nitrogen fixation
AZA90018		Rieske (2Fe-2S) iron-sulfur domain (5)	Nitrogen fixation
AZA90022		conserved hypothetical protein (6)	Nitrogen fixation
AZA90396		nitrogenase-associated protein (1)	Nitrogen fixation
AZA90402		iron-sulfur cluster assembly accessory protein (7)	Nitrogen fixation
AZA90634		conserved hypothetical protein (2)	Nitrogen fixation
AZA44366	glnK2	nitrogen regulatory protein P-II, GlnK2	Nitrogen metabolism
AZA53589	glnA	glutamine synthetase, type I	Nitrogen metabolism
AZA57320	rpoN	RNA polymerase factor sigma N	Nitrogen metabolism
AZA87992	ntrC	nitrogen assimilation regulatory protein	Nitrogen metabolism
AZA87994	ntrY	nitrogen regulation protein NtrY	Nitrogen metabolism
AZA87996	ntrX	nitrogen assimilation regulatory protein NtrX	Nitrogen metabolism
AZA88497	glnE	glutamine synthetase adenylyltransferase	Nitrogen metabolism
AZA89468	ntrY	nitrogen regulation protein NtrY	Nitrogen metabolism
AZA90389	amtB2	ammonium transporter	Nitrogen metabolism
AZA90638	gltD	glutamate synthase, small subunit	Nitrogen metabolism
AZA90639	glnK	nitrogen regulatory protein P-II, GlnK	Nitrogen metabolism
AZA90640	aat	aminotransferase	Nitrogen metabolism
AZA90641	ntrB	nitrogen regulation protein NtrB	Nitrogen metabolism
AZA90648	glnB	nitrogen regulatory protein P-II, GlnB	Nitrogen metabolism
AZA90654	gltB	glutamate synthase, large subunit	Nitrogen metabolism
AZA32850	phbB	acetoacetyl-CoA reductase	Poly-hydroxybutyrate biosynthesis
AZA72336	phbC	poly-beta-hydroxybutyrate polymerase	Poly-hydroxybutyrate biosynthesis
AZA88576	phbB	acetoacetyl-CoA reductase	Poly-hydroxybutyrate biosynthesis
AZA89083	phbA	acetyl-CoA acetyltransferase	Poly-hydroxybutyrate biosynthesis
AZA89855	phbA	beta-ketothiolase	Poly-hydroxybutyrate biosynthesis
AZA90629	phaZ	polyhydroxyalkanoate depolymerase	Poly-hydroxybutyrate degradation

AZA63403	luxR	transcriptional regulator, LuxR family	Quorum sensing
AZA90644	luxI	acyl-homoserine-lactone synthase	Quorum sensing
AZA25986	salB	beta-glucosidase	Salicin degradation
AZA26000	salA	beta-glucosidase	Salicin degradation
AZA87787		alpha-glucosidase	Sucrose degradation
AZA01854		TonB-dependent receptor	TonB-dependent transporter
AZA04400		TonB-dependent siderophore receptor	TonB-dependent transporter
AZA06033		TonB-dependent receptor	TonB-dependent transporter
AZA06440		TonB-dependent receptor	TonB-dependent transporter
AZA06565		TonB-dependent receptor	TonB-dependent transporter
AZA07052		TonB-dependent receptor	TonB-dependent transporter
AZA07492		TonB-dependent receptor	TonB-dependent transporter
AZA07658		TonB-dependent receptor	TonB-dependent transporter
AZA08067		TonB-dependent receptor	TonB-dependent transporter
AZA09133		TonB-dependent receptor	TonB-dependent transporter
AZA09279		TonB-dependent receptor	TonB-dependent transporter
AZA10816		TonB-dependent outer membrane receptor	TonB-dependent transporter
AZA14532		TonB-dependent siderophore receptor	TonB-dependent transporter
AZA15559		TonB-dependent receptor	TonB-dependent transporter
AZA15713		TonB-dependent receptor	TonB-dependent transporter
AZA15905		TonB-dependent receptor	TonB-dependent transporter
AZA16596		TonB-dependent receptor	TonB-dependent transporter
AZA16651		TonB-dependent receptor	TonB-dependent transporter
AZA16828		TonB-dependent receptor	TonB-dependent transporter
AZA17637		TonB-dependent receptor	TonB-dependent transporter
AZA18344		TonB-dependent receptor	TonB-dependent transporter
AZA21127		TonB-dependent receptor, vitamin B12/cobalamin transport	TonB-dependent transporter
AZA22051		TonB-dependent sideophore receptor	TonB-dependent transporter
AZA22446		TonB dependent receptor	TonB-dependent transporter
AZA22616		TonB-dependent receptor protein	TonB-dependent transporter
AZA22774		TonB-dependent receptor	TonB-dependent transporter
AZA24207		TonB dependent receptor	TonB-dependent transporter
AZA24508		TonB-dependent receptor	TonB-dependent transporter
AZA24558		TonB-dependent receptor	TonB-dependent transporter
AZA26015	cirA	TonB-dependent receptor	TonB-dependent transporter
AZA29067		TonB-dependent receptor	TonB-dependent transporter
AZA29136		TonB-dependent receptor	TonB-dependent transporter

AZA32595		TonB-dependent receptor	TonB-dependent transporter
AZA34449		TonB-dependent receptor	TonB-dependent transporter
AZA36912		TonB-dependent receptor	TonB-dependent transporter
AZA36929		TonB-dependent receptor	TonB-dependent transporter
AZA37846		TonB-dependent receptor	TonB-dependent transporter
AZA38746		tol biopolymer transport system	TonB-dependent transporter
AZA39144		TonB-dependent receptor	TonB-dependent transporter
AZA40366		TonB-dependent siderophore receptor	TonB-dependent transporter
AZA40462		TonB-dependent receptor, putative	TonB-dependent transporter
AZA41560		TonB-dependent receptor	TonB-dependent transporter
AZA42118		TonB-dependent receptor	TonB-dependent transporter
AZA42312		TonB-dependent receptor	TonB-dependent transporter
AZA43305		TonB-dependent receptor	TonB-dependent transporter
AZA43763		TonB-dependent outer membrane receptor protein	TonB-dependent transporter
AZA44478		TonB-dependent receptor	TonB-dependent transporter
AZA44838		TonB-dependent receptor	TonB-dependent transporter
AZA45743		TonB-dependent siderophore receptor	TonB-dependent transporter
AZA47317		TonB-dependent receptor	TonB-dependent transporter
AZA49407	iroN	TonB-dependent receptor	TonB-dependent transporter
AZA49939		TonB-dependent receptor	TonB-dependent transporter
AZA50366		TonB-dependent receptor	TonB-dependent transporter
AZA50524		TonB-dependent receptor	TonB-dependent transporter
AZA50533		TonB-dependent receptor	TonB-dependent transporter
AZA50691		TonB-dependent receptor	TonB-dependent transporter
AZA52119		TonB-dependent receptor	TonB-dependent transporter
AZA58525	fhuA	TonB-dependent receptor	TonB-dependent transporter
AZA58695		TonB-dependent receptor	TonB-dependent transporter
AZA59668		TonB-dependent siderophore receptor	TonB-dependent transporter
AZA60277		tonB dependent receptor	TonB-dependent transporter
AZA61044		TonB-dependent receptor	TonB-dependent transporter
AZA61526		TonB-dependent receptor Environmental Information Processing	TonB-dependent transporter
AZA61533		TonB-dependent receptor	TonB-dependent transporter
AZA64315		TonB-dependent receptor	TonB-dependent transporter
AZA65154		TonB-dependent receptor, plug	TonB-dependent transporter
AZA67151		TonB-dependent receptor, plug	TonB-dependent transporter
AZA68973		TonB-dependent receptor	TonB-dependent transporter
AZA69179		TonB-dependent receptor	TonB-dependent transporter

AZA70579		TonB-dependent receptor, putative	TonB-dependent transporter
AZA71591		TonB-dependent receptor, putative	TonB-dependent transporter
AZA72354		TonB-dependent receptor	TonB-dependent transporter
AZA73082		TonB-dependent receptor	TonB-dependent transporter
AZA73645		TonB-dependent receptor	TonB-dependent transporter
AZA74103		TonB-dependent receptor	TonB-dependent transporter
AZA76890		TonB-dependent receptor	TonB-dependent transporter
AZA79474		TonB-dependent receptor	TonB-dependent transporter
AZA79883		TonB-dependent receptor	TonB-dependent transporter
AZA80916		TonB-dependent receptor	TonB-dependent transporter
AZA80968		TonB-dependent receptor	TonB-dependent transporter
AZA81511	exbD	Biopolymer transport protein ExbD	TonB-dependent transporter
AZA87233		TonB-dependent receptor	TonB-dependent transporter
AZA87290		TonB-dependent receptor	TonB-dependent transporter
AZA87507	exbD	TonB system transport protein ExbD type-1	TonB-dependent transporter
AZA87509	TonB	TonB protein	TonB-dependent transporter
AZA87760		TonB-dependent receptor	TonB-dependent transporter
AZA87789		TonB-dependent receptor	TonB-dependent transporter
AZA87908		TonB-dependent receptor	TonB-dependent transporter
AZA87957		TonB-dependent receptor	TonB-dependent transporter
AZA88020		TonB-dependent receptor plug domain protein	TonB-dependent transporter
AZA88034		TonB-dependent receptor	TonB-dependent transporter
AZA88052		TonB dependent receptor	TonB-dependent transporter
AZA88204		TonB-dependent receptor	TonB-dependent transporter
AZA88227		TonB-dependent receptor	TonB-dependent transporter
AZA88329		TonB-dependent receptor	TonB-dependent transporter
AZA88359		TonB-dependent receptor	TonB-dependent transporter
AZA88461	tolQ	biopolymer transport protein	TonB-dependent transporter
AZA88767		TonB-dependent receptor	TonB-dependent transporter
AZA88861		TonB-dependent receptor	TonB-dependent transporter
AZA88894		TonB-dependent receptor	TonB-dependent transporter
AZA89263	bfrD	TonB-dependent receptor bfrD	TonB-dependent transporter
AZA89405		TonB-dependent receptor	TonB-dependent transporter
AZA89496		TonB-dependent receptor	TonB-dependent transporter
AZA89590		TonB-dependent receptor	TonB-dependent transporter
AZA89602		ExbD	TonB-dependent transporter
AZA89921		TonB-dependent receptor domain protein	TonB-dependent transporter

AZA90193		TonB-dependent siderophore receptor	TonB-dependent transporter	
AZA90405		biopolymer transport protein exbB	TonB-dependent transporter	
AZA90406		biopolymer transport ExbD protein	TonB-dependent transporter	
AZA90407		biopolymer transport exbD1 protein	TonB-dependent transporter	
AZA90559		TonB-dependent receptor yncD	TonB-dependent transporter	
AZA87477	ureC	urease, subunit alpha UreC	Urea metabolism	
AZA87480	ureA	urease, gamma subunit UreA	Urea metabolism	
AZA87481	ureD	Urease accessory protein UreD	Urea metabolism	
AZA87506	ureF	urease accessory protein UreF	Urea metabolism	
AZA87516	ureE	urease accessory-like protein UreE	Urea metabolism	
AZA87589	ureG	urease accessory protein UreG	Urea metabolism	
AZA90633	ureB	urease, beta subunit	Urea metabolism	
AZA15221	glnD	P-II uridylyltransferase	Nitrogen metabolism	
AZA15920		nitrilase	Phytormone production	
AZA87569	cheA2	CheA signal transduction histidine kinase	Chemotaxis	CheA2
AZA90328		CheA signal transduction histidine kinase	Chemotaxis	CheA3
AZA90451		CheA Signal transduction histidine Kinase (STHK)	Chemotaxis	CheA (C-terminus homolog to CheY)
AZA87483	cheAY	CheAY	Chemotaxis	CheA (C-terminus homolog to CheY)
AZA87567	cheC	CheC, inhibitor of MCP methylation	Chemotaxis	
AZA87458	cheL	chemotactic signal-response protein CheL	Chemotaxis	CheL
AZA87492	cheR	chemotaxis methyltransferase CheR	Chemotaxis	CheR
AZA80979	motA	chemotaxis MotA protein	Chemotaxis	
AZA90662	motA	chemotaxis motA protein	Chemotaxis	
AZA51771	motB	chemotaxis MotB protein	Chemotaxis	
AZA85137	motB	chemotaxis MotB protein	Chemotaxis	
AZA88725	motB	chemotaxis MotB protein	Chemotaxis	
AZA90369		chemotaxis MotB protein	Chemotaxis	
AZA87637	cheY	chemotaxis protein	Chemotaxis	CheY
AZA87485	cheY	chemotaxis protein CheY	Chemotaxis	CheY
AZA88408	cheY	chemotaxis protein cheY	Chemotaxis	CheY
AZA49142	cheYI	chemotaxis protein CheYI	Chemotaxis	
AZA11540	cheZ	chemotaxis protein CheZ	Chemotaxis	CheZ
AZA90691		chemotaxis two-component response regulator protein-glutamate methylesterase	Chemotaxis	CheB2
AZA88501		chemotaxis-specific methylesterase	Chemotaxis	CheB2
AZA87484	cheW	CheW protein	Chemotaxis	CheW
AZA90470	cheW	CheW protein	Chemotaxis	CheW2
AZA76052	cheW	CheW protein	Chemotaxis	CheW

AZA90452		CheW protein	Chemotaxis	CheW2
AZA90497		CheY-like receiver	Chemotaxis	CheYIII
AZA89006		CheY-like response regulator	Chemotaxis	
AZA90323		response regulator receiver (CheY-like) modulated CheB methylesterase	Chemotaxis	CheB3
AZA90319	cheB	response regulator receiver modulated CheB methylesterase	Chemotaxis	CheB2
AZA85614	cheR3	MCP methyltransferase, CheR-type	Chemotaxis	CheR3
AZA90318		MCP methyltransferase, CheR-type	Chemotaxis	CheR2
AZA90685		MCP methyltransferase, CheR-type	Chemotaxis	CheR2
AZA90217		MCP methyltransferase/methylesterase, CheR/CheB with PAS/PAC sensor	Chemotaxis	
AZA89703	fliH	flagellar assembly protein FliH	Flagellum	
AZA87530	fliO	flagellar assembly protein FliO	Flagellum	
AZA23600	fliX	flagellar assembly regulator FliX	Flagellum	
AZA87459	flgG	flagellar basal body distal rod protein - FlgG	Flagellum	
AZA88657	flgA	flagellar basal body P-ring biosynthesis protein FlgA	Flagellum	
AZA87475	flgC	flagellar basal body rod protein	Flagellum	
AZA87537	fliL	flagellar basal body-associated protein FliL	Flagellum	
AZA87474	flgB	flagellar basal-body rod protein FlgB	Flagellum	
AZA23566	flgF	flagellar basal-body rod protein FlgF	Flagellum	
AZA82102	flhA	flagellar biosynthesis	Flagellum	
AZA87469	flhB	flagellar biosynthesis protein FlhB	Flagellum	
AZA87471	fliQ	flagellar biosynthesis protein FliQ	Flagellum	
AZA60822	fliR	flagellar biosynthesis protein FliR	Flagellum	
AZA87528	flhB	flagellar biosynthetic protein FlhB	Flagellum	
AZA87468	fliP	flagellar biosynthetic protein fliP	Flagellum	
AZA87463	flhF	flagellar GTP-binding protein	Flagellum	
AZA70724	flgD	flagellar hook capping protein FlgD	Flagellum	
AZA11648	flgE	flagellar hook protein FlgE	Flagellum	
AZA31666	flgE	flagellar hook protein FlgE precursor	Flagellum	
AZA78244	flaN	flagellar hook-associated protein	Flagellum	
AZA88315	flgK	flagellar hook-associated protein 1	Flagellum	
AZA87546	flgL	flagellar hook-associated protein FlgL	Flagellum	
AZA88242	fliE1	flagellar hook-basal body complex protein fliE 1	Flagellum	
AZA88601	fliK	flagellar hook-length control protein	Flagellum	
AZA87578	flgH	flagellar L-ring protein FlgH	Flagellum	
AZA42701	fliG	flagellar motor switch protein FliG	Flagellum	
AZA90379	fliM	flagellar motor switch protein FliM	Flagellum	
AZA87457	motA	flagellar polar motor protein MotA	Flagellum	

AZA90583	flgI	flagellar P-ring protein FlgI	Flagellum	
AZA89798	flbT	flagellar protein	Flagellum	
AZA89796		flagellar synthesis regulator, FleN	Flagellum	
AZA90646	flaA	flagellin	Flagellum	
AZA36138		flagellin domain-containing protein	Flagellum	
AZA89806		flagellin domain-containing protein	Flagellum	
AZA87550	fliS	flagellin-specific chaperone FliS	Flagellum	
AZA90173	fliI	flagellum-specific ATP synthase	Flagellum	
AZA87460	fliN	polar flagellar motor switch protein FliN	Flagellum	
AZA78256	fliD	putative flagellar hook-associated protein 2 - FliD	Flagellum	
AZA76775		methyl-accepting chemotaxis protein	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA27777		methyl-accepting chemotaxis sensory transducer with PAS/Pac sensor Aer-like	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA84986		methyl-accepting chemotaxis sensory transducer with PAS/Pac sensor Aer-like	Chemotaxis	
AZA90554		methyl-accepting chemotaxis	Chemotaxis	mcp I, serine sensor receptor
AZA88794	mcp	methyl-accepting chemotaxis protein	Chemotaxis	
AZA89515	mcp	methyl-accepting chemotaxis protein	Chemotaxis	
AZA89528	mcp	methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA89608	mcp	methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA89758	mcp	methyl-accepting chemotaxis protein	Chemotaxis	aerotaxis sensor receptor
AZA90070	mcp	methyl-accepting chemotaxis protein	Chemotaxis	aerotaxis sensor receptor
AZA02033		methyl-accepting chemotaxis protein	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA23142		methyl-accepting chemotaxis protein	Chemotaxis	
AZA27812		methyl-accepting chemotaxis protein	Chemotaxis	mcp II, aspartate chemoreceptor protein
AZA28518		methyl-accepting chemotaxis protein	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA41421		methyl-accepting chemotaxis protein	Chemotaxis	mcp II
AZA63880		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA80173		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA81161		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA83437		methyl-accepting chemotaxis protein	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA88489		methyl-accepting chemotaxis protein	Chemotaxis	mcp II, aspartate chemoreceptor protein
AZA88686		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA88870		methyl-accepting chemotaxis protein	Chemotaxis	mcp II, aspartate chemoreceptor protein
AZA89115		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA89563		methyl-accepting chemotaxis protein	Chemotaxis	
AZA89676		methyl-accepting chemotaxis protein	Chemotaxis	mcp III, ribose and galactose sensor receptor
AZA90000		methyl-accepting chemotaxis protein	Chemotaxis	
AZA90197		methyl-accepting chemotaxis protein	Chemotaxis	mcp IV, dipeptide chemoreceptor protein

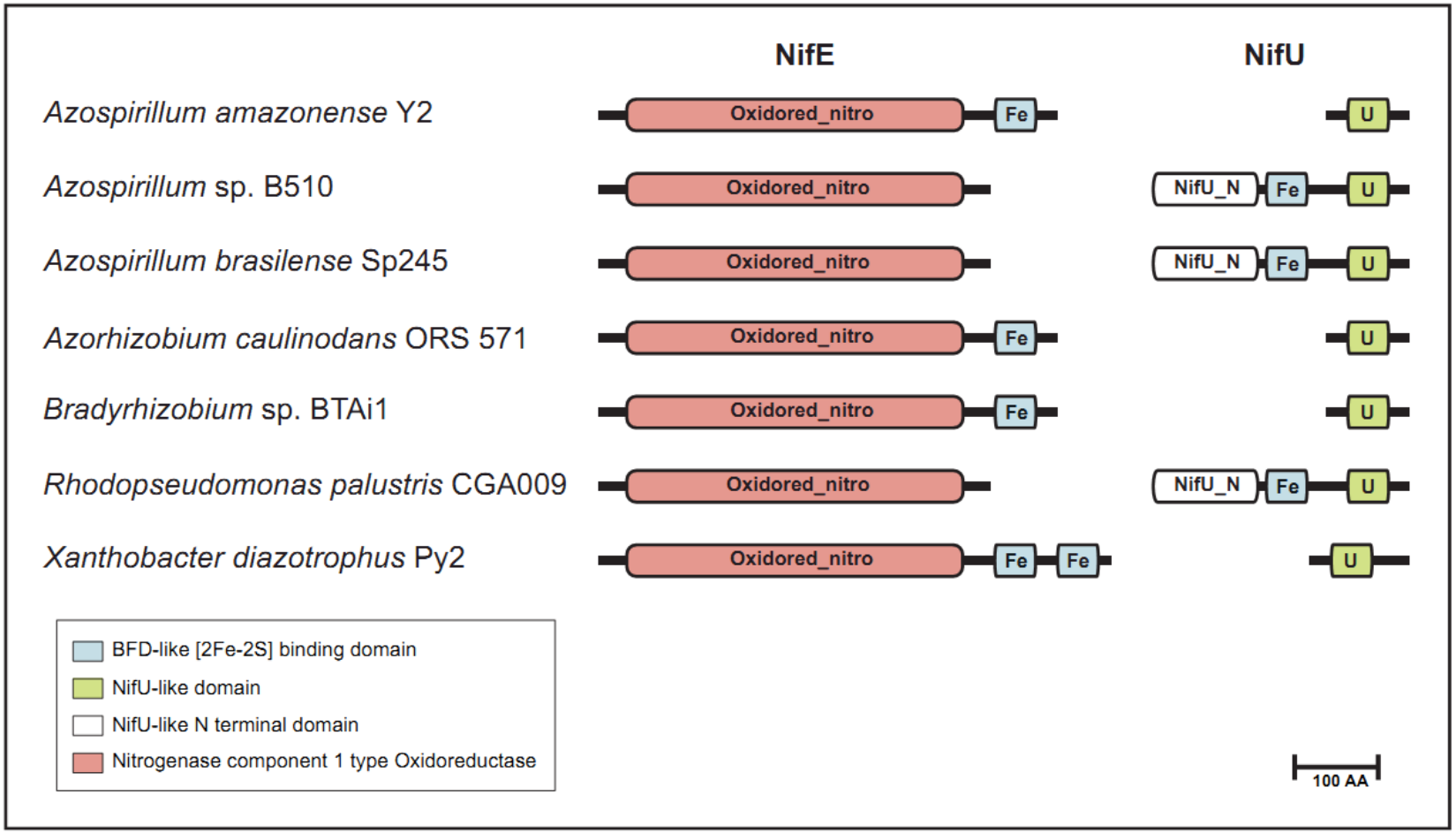
AZA90556		methyl-accepting chemotaxis protein	Chemotaxis	mcp I, serine sensor receptor
AZA19614	mcp2	methyl-accepting chemotaxis protein 2	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA88383	mcp4	methyl-accepting chemotaxis protein 4	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA90386	mcp	methyl-accepting chemotaxis protein McpB	Chemotaxis	aerotaxis sensor receptor
AZA88564	mcpB	methyl-accepting chemotaxis protein McpB	Chemotaxis	mcp IV, dipeptide chemoreceptor protein
AZA88486	mcp	methyl-accepting chemotaxis receptor/sensory transducer	Chemotaxis	
AZA12638		methyl-accepting chemotaxis sensory transducer	Chemotaxis	
AZA21719		methyl-accepting chemotaxis sensory transducer	Chemotaxis	mcp II, aspartate chemoreceptor protein
AZA37653		methyl-accepting chemotaxis sensory transducer	Chemotaxis	
AZA86491		methyl-accepting chemotaxis sensory transducer	Chemotaxis	mcp I, serine sensor receptor
AZA88057		methyl-accepting chemotaxis sensory transducer	Chemotaxis	mcp I, serine sensor receptor
AZA90327		methyl-accepting chemotaxis sensory transducer	Chemotaxis	
AZA26184		methyl-accepting chemotaxis sensory transducer with Pas/Pac sensor	Chemotaxis	aerotaxis sensor receptor
AZA88611	mcp	putative methyl-accepting chemotaxis protein	Chemotaxis	mcp II, aspartate chemoreceptor protein
AZA90004		methyl-accepting chemotaxis sensory transducer	Chemotaxis	mcp II, aspartate chemoreceptor protein

INCOMPLETE ORFS

Name	Product	Category	Sequence 1	Sequence 2
fixI	cation transport ATPase	Microaerophilic respiration	MPLDSLTTHCRHCGQPLLGAATERGAFCCQGCAGAYTLI HDLGLDAYYDRRPAEAAAPAVADDGDGADPLLGLP	VSLSPASAVDVTQTAADVVFQGDGLAAVAETLAVARRARAL VRQNLAFTLLYNLCAVPVAIAGFATPFVAAAAMSGSSLAVM ANALRLGSKKPRDGGYR
pfkA	6-phosphofructokinase	Glycolysis	MSNGKRIGILTSGGDCAGLNAVIRAVVHRAVTTYSWQVL GIEEGTHGLLDRPVRHRVLLPAHVDGNILRQGGTVLGTT NKGDPFAYPMPDGTKKDRSQEIISGCKELGLDALIGIGG DGSLAILRRLAQEGGLNLVGIPKTI	VMVLEVMGRDAGHIALAAGIAGGADVILLPEIPWTVDGVAN RIRQIRAQGRNFALVVVSEAAALPGGGKARQEFTDGQKRYG GIGSYIGHLIAEATGAETRVTVLGHVQRGGSPNPRDRLIAS AFGVHAVDLIAHGKYDRMVAWSNRKVIDVPIADAIKAYAAV EVDGTLVRTARGLGIYVGE
amtB	ammonium transporter	Nitrogen metabolism	TISGGDTAWMLMSTALVLMMTIPGLALFYGGMVRKFNVL AMVMQSFTITCLVTVIWALVGYSLSFTAASDAADAKEFI GGFSRVLLAGMDPTGTHALAPTIPEPVFMMYQMTFAIIT PSLITGAFADRMKFSSMLVFIALWSLLVYAPVAHTVWHP NGFMAKLGVLDFAGGTVVHINAGIAGLVACLVIGKRKGY PNADFSPHNLVLSLIGASLLWVGWFGFNAGSALTAGPRA GMAMTVTQVATGAAALAWVFVEWALKGKPSVLGIISGAV AGLVAITPAAGFVDTVGAFAIGIAAGVICFFTATSMKRA LGYDDSLDAFGVHGIGGIVGAILTGVFAKMSIANSEGAF AATLKDAPDTKLGLIEGNVGQVLTQLEGVAFTVVWCAVI TFILLKIIDLVIGLRVDEDIERDGLDIALHGESIH	

nasA	nitrate reductase	Nitrate/nitrite metabolism	MARDAEAGALTGGVRSTCPYCGVGCGVVAQADGRVVGDT DHPANRGRLCSKGSALAETLKDGERLTRPLIDGRAADWT QALDLVAQRFKATIAEHGPDSVAFYVSGQCLTEDYYVAN KLMKGFIGSGNIDTNSRLCMASSVAGHNRAFGADVVPGT YEDLELADLVVLVGSNSAWCHPVLHQRLLAAKAKRPLRI VAIDPRRTATCEDADLHLAVRPGADVALFNGLLAHLADQ DQLDHAWMAAHAEGFEAALAAARQTADQTAALTGLDPAD LARFYQWFASTPRVVTVYSQGVNQSSVGTDKVNAIINCH LATGRIGKPGCGPFSITGQPNAMGGREVGGLANQLAAHM RFDQPEAIDRVRRFWRAPNLATKPGLKAVEMFDAVLDGR IKALWIIATNPADSLPRADQVREALRACPFVVVSDCWPN DTTALADVVLPAAGWGERDGTVTNSERVISRQRPFRPAP GEARADWRV	TDAGRGQLRVAILVDGRLEACLFLAAGDGAQLPPRDVLAAL LGDRVTDEARPSLLSGRAPGGSGAMGVKSTGRTVCACFSVG LVTIRDAIVAKRLTSAEEIGACLKAGTNCGSCIPELKEILR DVHASA
nasD	nitrite reductase large subunit	Nitrate/nitrite metabolism	MPPRQKLVVVGNGMAGMRTVDELLKRNPLKYDITVFGAE PHVNYDRIMLSSVLAGEKDLEQIVINPRSWYEENGIELI TGDAVTAIDTQARTVTSASGRVVAYDKVLLATGSRPIAP PVPGLDLPGVCAFRDIADVNTMIAASQTHKRAIVIGGGL LGLEAANGLLRRGMQVAVVHLMGTLMERQLDKAAAELLQ RELDERGMNFFTNGQTEEIFGEGRVQGVRLADGREIPGD LVVLAIGIRPNIDLARAAGLAINRGIEVGDDMRTSVPDI FAVGECVEHRGKTYGLVAPLWDMAKVCADHMAREDATSD YVPAAMATRLKVTGIDVFSAGDFIGDETTEEVVFRDAAR NTYKRLVLRDDRLVGAVLYGDAQDGGWYFQLLREDKPLG ELRD	KGKIMGAIAEHGLTTLEGVRARTKASASCGSCTPQVEQLLA LSLGDSYAKPTGEKPLCKCTTHSHDQVRRAIVDQELRTMPA VMSTLGWSTPDGCHVCRPALNFYLLVAWPGEYQDDQQSRFV NERNHANIQKDGTYSVVPRMWGGLTTADELRAIADVVDKYA IPTVKVTGGQRIDLLGVTREQLPHVWKDLNDAGMVSGHAYG KALRTVKTCVGSEWCRFGTQDSTGMGVKLERMTWGSWHAHK VKLAVSGCPRNCAEATIKDFGVVAVDSGWELHVGGNGGIHV RVTDLLCKVTTEEEVLEYCGAYLQLYREESRYLERTAPWIE RVGLAHVKQRIVDDEAGRKALYAKFLYAQKFSQDDPWAERA PGKAAAAPFRTLATVE

fliF	flagellar MS-ring protein	Flagellum	LENLLQTLRNLGPVRLGAIGGVALLILALFGFLLFNGSS PNMALLYSDLSPSDGGAIVQQLDQMQPKVPYQVSPDSTR IEVPADQVGRIRMLMAQQGLPTGGNVGYEIFNQAESLGT TSFMQSVQQLRALEGELSRTVNTLAPVQQSRIHLVLPKR ELFSRETQKATASVVLRLRPGQQLKKEQVASVQHLIAAS VPGLQPDMVSVVDDKGNLLARGMGSDSKEAMMATAEEKR LAYQQRLTDKVEEIVGRTVGMGKVRAEVTVDMDFDHITT NSEIFDPDSQVVRSTETNNENSEDINKEGQDPVTVANNL PSANAENTTNGSSSSKTTKTNERINYEISKTVKVHEREA GQVRRLSVAVLVDGNYVPDDKGVAQYQPRSDDELQKLGS LVKSSIGFDGSRGDTVDVVSMKFATPEGELEAKEETLFG LPKQDVFRIAETIVLAIVAILVILLVIRPLVARALDRTP QLDEEPDLLSDQSGVP
flbD	transcriptional regulatory protein FlbD (LafK)	Flagellum	MRLLIVGTLEGYITAAGKIALQKGAKVAHTDSIGGALNA LRGGQGADLVMIDVKLDVAMLIEALKTERITVPVVACGV ATDAQAAVKAIRAGAKEYIPLPPDAQLIAAVLEAVAEES HAIVSRDPSMGSVLRLADQIAPSDASVLITGESGTGKEL MARYIHRKSRRANAVFVSVNCAAIPENLLESELFGHEKG AFTGAVARRVGKFEEANGGTLLLDEISEMDIRLQAKLLR AIQEREIDRVGGTQPVKVDIRVLATSNRFLEEEVRAGRF REDLFFRLNVVNLALPPLRERPADIAILSQHFATKYSEA NGLPERRVSPEAMAMLQAHRFRGNVRELENTMHRAVLLS RGTEIGTDAILLTGAEPAAPAQGYSAVAPVAA



Additional file 2. Supplementary figure 1. Domain composition of the NifE and NifU proteins among bacteria from the orders Rhodospirillales and Rhizobiales. The oblong boxes represent protein domains which are colored according to the description in the legend.

Additional file 3. Supplementary table 2. Putative sigma N and NifA binding sites occurring upstream of some *nif* and *fix* genes.

Gene	Accession number	Sigma 54 binding sites	Distance from ATG (bp)	Reference
nifH	AZA90011	CACT**TGGCAC**ACCGG**TTGC**TGAAA	157	
iron-sulfur cluster assembly accessory protein gene (7)	AZA90402	GCCG**TGGCACG**CCCC**TTGC**GGAGG	31	
fixA	AZA90029	CAGT**TGGCA**T**G**GCCC**TTGC**CAATT	41	
nifB	AZA65617	GACG**TGGCACG**GCCT**TTGC**TGAGT	98	
Consensus		mrNrY**TGGCACG**NNNN**TTGC**WNNw		Barrios *et al.*, 1999

Gene	Accession number	NifA binding sites	Distance from ATG (bp)	Reference
nifH	AZA90011	**TGT**TCCGTTTGCG**ACA**	274	
nifH	AZA90011	**TGT**CGGGTTCGGC**ACA**	245	
iron-sulfur cluster assembly accessory protein gene (7)	AZA90402	**TGT**CGCAAACCCG**ACA**	113	
fixA	AZA90029	**TGT**CGCCTATCCG**ACA**	125	
nifB	AZA65617	**TGT**CGATAACCGC**ACA**	175	
Consensus		**TGT**NNNNNNNNNN**ACA**		Buck *et al.*, 1986

References

Compilation and analysis of σ54-dependent promoter sequences

Nucleic Acids Research, Vol. 27, No. 22. (1 November 1999), pp. 4305-4313.

Humberto Barrios, Brenda Valderrama, Enrique Morett

Upstream activator sequences are present in the promoters of nitrogen fixation genes

Nature, Vol. 320, No. 6060. (27 March 1986), pp. 374-378.

Martin Buck, Stephen Miller, Martin Drummond, Ray Dixon

Additional file 4. Supplementary table 3. Accession numbers of the sequences utilized in the phylogenetic reconstructions.

16S

Species	16S accession number or sequence	Taxonomic group
Azorhizobium caulinodans ORS 571	D11342	Rhizobiales (Alphaproteobacteria)
Azorhizobium doebereinerae BR5401	AF391130	Rhizobiales (Alphaproteobacteria)
Azospirillum amazonense Y2	AGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCATGCCTAACACAT GCAAGTCGAACGAAGGCTTCGGCCTTAGTGGCGCACGGGTGAGTAACGC GTGGGAACCTGCCCTTTGGTTCGGAATAACTCCGGGAAACTGGAGCTAA TACCGGATGAGCCGGTTGGCCGTGGAGGCTGATCGGGAAAGATTTATCG CCGAAGGAGGGGCCCGCGTCCGATTAGGTAGTTGGTGAGGTAACGGCTC ACCAAGCCGACGATCGGTAGCTGGTCTGAGAGGATGATCAGCCACACTG GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATAT TGGACAATGGGGGCAACCCTGATCCAGCAATGCCGCGTGAGTGATGAAG GCCTTAGGGTTGTAAAGCTCTTTCGCACGTGACGATGATGACGGTAACG TGAGAAGAAGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAA GGGGGCTAGCGTTGTTCGGAATTACTGGGCGTAAAGGGCGCGTAGGCGG CTGTTCAAGTCAGGCGTGAAATCCCCGGGCTTAACCTGGGAATTGCGTT TGAGACTGAGCGGCTTGAGTTCGGGAGAGGTGAGTGGAATTCCCAGTGT AGAGGTGAAATTCGTAGATATTGGGAAGAACACCGGTGGCGAAGGCGGC TCACTGGACCGATACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACA GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGACGT TGGGGTCCTTAGGACTTCGGTGTCGCAGCTAACGCATTAAGCACTCCGC CTGGGGAGTACGGCCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCC CGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGCAGAACC TTACCAGCTCTTGACATCTACTGACCGGTGCGGAGACGCACTTTCCACT TCGGTGGCGGTAAGACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGT GAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCCATCTTCAGTTG CCATCATTAGGTTGGGCACTCTGGAGAAACTGCCGGTGACAAGCCGGAG GAAGGCGGGGATGACGTCAAGTCCTCATGGCCCTTATGGGCTGGGCTAC ACACGTGCTACAATGGTGGTGACAGTGGGCAGCGACCTCGTGAGGGGGA GCTAATCTCCAAAAGCCATCTCAGTTCGGATTGTACTCTGCAACTCGAG TGCATGAAGTTGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGA ATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGG CTTTACCCGAAGCCGGTGCGCTAACCGCAAGGAGGCAGCCGACCACGGT ACGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACC TGCGGCTGGATCACCT	Rhodospirillales (Alphaproteobacteria)

Azospirillum brasilense Sp245	AGCGGCCCTGTCAAGTGGATGCGGTCTTCGGACCGGTGTCCAGCTTGAA CCTGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCATGCCTAAC ACATGCAAGTCGAACGAAGGCTTCGGCCTTAGTGGCGCACGGGTGAGTA ACACGTGGGAACCTGCCTTTCGGTTCGGGATAACGTCTGGAAACGGACG CTAACACCGGATACGTCCTTCGGGAGAAAGTTTACGCCGAGAGAGGGGC CCGCGTCCGATTAGGTAGTTGGTGGGGTAATGGCCCACCAAGCCGACGA TCGGTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACG GCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGGG CAACCCTGATCCAGCAATGCCGCGTGAGTGATGAAGGCCTTAGGGTTGT AAAGCTCTTTCGCACGCGACGATGATGACGGTAGCGTGAGAAGAAGCCC CGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCGAGCGTT GTTCGGAATTACTGGGCGTAAAGGGCGCGTAGGCGGCCTGTTTAGTCAG AAGTGAAAGCCCCGGGCTTAACCTGGGAACGGCTTTTGATACTGGCAGG CTTGAGTTCCGGAGAGGATGGTGGAATTCCCAGTGTAGAGGTGAAATTC GTAGATATTGGGAAGAACACCGGTGGCGAAGGCGGCCATCTGGACGGAC ACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCC TGGTAGTCCACGCCGTAAACGATGAATGCTAGACGCTGGGGTGCATGCA CTTCGGTGTCGCCGCTAACGCATTAAGCATTCCGCCTGGGGAGTACGGC CGCAAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGG AGCATGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAACCCTTGA CATGTCCACCACCGGCTCCAGAGATGGAGCTTTCAGTTCGGCTGGGTGG AACACAGGTGCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGG TTAAGTCCCGCAACGAGCGCAACCCCTACCGCCAGTTGCCATCATTCAG TTGGGCACTCTGGTGGAACTGCCGGTGACAAGCCGGAGGAAGGCGGGGA TGACGTCAAGTCCTCATGGCCCTTATGGGTTGGGCTACACACGTGCTAC AATGGCGGTGACAGTGGGATGCGAAGTCGCAAGATGGAGCCAATCCCCA AAAGCCGTCTCAGTTCGGATTGCACTCTGCAACTCGGGTGCATGAAGTT GGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCG GGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGGCTTTACCCGAA GGTGGTGCGCTAACCCGCAAGGG	Rhodospirillales (Alphaproteobacteria)
Azospirillum brasilense Sp7	AY324110	Rhodospirillales (Alphaproteobacteria)
Azospirillum canadense DS2	DQ393891	Rhodospirillales (Alphaproteobacteria)
Azospirillum doebereinerae 63f	AJ238567	Rhodospirillales (Alphaproteobacteria)
Azospirillum halopraeferens DSM 3675	X79731	Rhodospirillales (Alphaproteobacteria)
Azospirillum irakense Z29583	Z29583	Rhodospirillales (Alphaproteobacteria)
Azospirillum lipoferum (ncimb 11861)	Z29619	Rhodospirillales (Alphaproteobacteria)
Azospirillum lipoferum DSM 1691	GU256441	Rhodospirillales (Alphaproteobacteria)
Azospirillum melinis TMCY 0552	DQ022958	Rhodospirillales (Alphaproteobacteria)
Azospirillum melinis LMG 24250	GU256442	Rhodospirillales (Alphaproteobacteria)
Azospirillum oryzae COC8	AB185396	Rhodospirillales (Alphaproteobacteria)
Azospirillum picis IMMIB TAR-3	AM922283	Rhodospirillales (Alphaproteobacteria)
Azospirillum rugosum IMMIB	AM419042	Rhodospirillales (Alphaproteobacteria)
Azospirillum sp. B510	AB049111	Rhodospirillales (Alphaproteobacteria)

Azospirillum zeae N7	DQ682470	Rhodospirillales (Alphaproteobacteria)
Bradyrhizobium sp. BTAi1	CP000494	Rhizobiales (Alphaproteobacteria)
Cyanothece sp. ATCC 51142	CP000806	Cyanobacteria
Magnetospirillum gryphiswaldense MSR-1	Y10109	Rhodospirillales (Alphaproteobacteria)
Magnetospirillum magnetotacticum DSM 3856	Y10110	Rhodospirillales (Alphaproteobacteria)
Mesorhizobium loti MAFF303099	BA000012	Rhizobiales (Alphaproteobacteria)
Nostoc azollae' 0708	CP002059	Cyanobacteria
Rhizobium etli CFN 42	NR_029184	Rhizobiales (Alphaproteobacteria)
Rhodocista pekingensis 3-p	AF523824	Rhodospirillales (Alphaproteobacteria)
Rhodopseudomonas faecalis gc	AF123085	Rhizobiales (Alphaproteobacteria)
Rhodopseudomonas palustris CGA009	BX572608	Rhizobiales (Alphaproteobacteria)
Rhodopseudomonas palustris ATCC 17001	L11664	Rhizobiales (Alphaproteobacteria)
Rhodopseudomonas rhenobacensis Klemme	AB087719	Rhizobiales (Alphaproteobacteria)
Rhodospirillum centenum SW	CP000613	Rhodospirillales (Alphaproteobacteria)
Rhodospirillum oryzae JA318	AM901295	Rhodospirillales (Alphaproteobacteria)
Rhodospirillum rubrum ATCC 11170	CP000230	Rhodospirillales (Alphaproteobacteria)
Rhodospirillum sulfurexigens JA143	AM710622	Rhodospirillales (Alphaproteobacteria)
Sinorhizobium meliloti 1021	AL591688	Rhizobiales (Alphaproteobacteria)
Xanthobacter agilis SA35	X94198	Rhizobiales (Alphaproteobacteria)
Xanthobacter aminoxidans 14a	AF399969	Rhizobiales (Alphaproteobacteria)
Xanthobacter autotrophicus 7c	X94201	Rhizobiales (Alphaproteobacteria)
Xanthobacter autotrophicus Py2	X94199	Rhizobiales (Alphaproteobacteria)
Xanthobacter flavus 301	X94199	Rhizobiales (Alphaproteobacteria)
Xanthobacter tagetidis TagT2C	X99469	Rhizobiales (Alphaproteobacteria)
Xanthobacter viscosus 7d	AF399970	Rhizobiales (Alphaproteobacteria)

CbbSL

Species	CbbS accession number	CbbL accession number	Form	Taxonomic group
Acidiphilium cryptum JF-5	YP_001233963	YP_001233962	IC	Rhodospirillales (Alphaproteobacteria)
Acidithiobacillus ferrooxidans	AAD30509	AAD30508	IA	Gammaproteobacteria
Acidithiobacillus ferrooxidans	AAK00292	Q07087	IA	Gammaproteobacteria
Azospirillum amazonense Y2	AZA90631	AZA80325	IC	Rhodospirillales (Alphaproteobacteria)
Bradyrhizobium japonicum USDA 110	NP_769226	NP_769225	IC	Rhizobiales (Alphaproteobacteria)
Bradyrhizobium sp. BTAi1	YP_001242212	YP_001242213	IC	Rhizobiales (Alphaproteobacteria)
Bradyrhizobium sp. BTAi1	YP_001236644	YP_001236643	IC	Rhizobiales (Alphaproteobacteria)
Bradyrhizobium sp. BTAi1	YP_001238691	YP_001238690	IA	Rhizobiales (Alphaproteobacteria)
Burkholderia xenovorans LB400	YP_552889	YP_552888	IC	Betaproteobacteria

Cupriavidus necator (*Ralstonia eutropha*)	P09658	P0C2C2	IC	Betaproteobacteria
Nitrosospira multiformis ATCC25196	YP_411384	YP_411385	IC	Betaproteobacteria
Paracoccus denitrificans P1222	ZP_00629447	ZP_00629446	IC	Rhodobacterales (Alphaproteobacteria)
Rhodobacter sphaeroides 2.4.1	YP_354362	YP_354363	IC	Rhodobacterales (Alphaproteobacteria)
Rhodobacter sphaeroides ATCC 17025	YP_001168903	YP_001168904	IC	Rhodobacterales (Alphaproteobacteria)
Rhodopseudomonas palustris CGA009	NP_946906	NP_946905	IC	Rhizobiales (Alphaproteobacteria)
Rhodospirillum centenum SW	ACJ01400	ACJ01401	IC	Rhodospirillales (Alphaproteobacteria)
Rhodospirillum centenum SW	ACI97688	ACI97689	IA	Rhodospirillales (Alphaproteobacteria)
Sinorhizobium meliloti 1021	NP_436730	NP_436731	IC	Rhizobiales (Alphaproteobacteria)
Xanthobacter autotrophicus Py2	YP_001416819	YP_001416820	IC	Rhizobiales (Alphaproteobacteria)
Xanthobacter flavus	P23012	P23011	IC	Rhizobiales (Alphaproteobacteria)

NifH

Species	NifH accession number or sequence	Taxonomic group
Azorhizobium caulinodans ORS 571	AP009384	Rhizobiales (Alphaproteobacteria)
Azospirillum amazonense Y2	AZA90011	Rhodospirillales (Alphaproteobacteria)
Azospirillum brasilense Sp245	ATGTCTTTGCGCCAGATTGCGTTCTACGGTAAGGGCGGTATCGGAAAGT CCACCACCTCCCAGAACACCCTGGCCGCGCTGGTCGAGCTGGATCAGAA GATCCTGATCGTCGGCTGCGATCCGAAGGCCGACTCGACCCGCCTGATC CTGCACGCCAAGGCGCAGGACACCGTGCTGCACCTCGCCGCCGAAGCCG GCTCGGTCGAGGATCTGGAACTCGAGGACGTTCTCAAGATCGGCTACAA GGGCATCAAGTGCGTCGAGTCCGGCGGTCCGGAGCCGGGGGTCGGCTGC GCCGGCCGCGGCGTGATCACCTCGATCAACTTCCTCGAAGAGAACGGCG CCTACGACGACGTGGACTACGTCTCCTACGACGTGCTGGGCGACGTGGT GTGCGGCGGTTTCGCCATGCCCATCCGCGAGAACAAGGCCCAGGAAATC TACATCGTCATGTCCGGTGAGATGATGGCGCTCTACGCCGCCAACAACA TCGCCAAGGGCATTCTGAAGTACGCGCACAGCGGCGGCGTGCGCCTCGG CGGCCTGATCTGCAACGAGCGCCAGACCGACAAGGAAATCGACCTCGCC TCGGCCCTGGCCGCCCGCCTCGGCACCCAGCTCATCCACTTCGTGCCGC GCGACAACATCGTGCAGCACGCCGAGCTGCGCCGCATGACCGTGATCGA GTACGCGCCGGACAGCCAGCAGGCCCAGGAATACCGCCAGCTCGCCAAC AAGGTCCACGCGAACAAGGGCAAGGGCACCATCCCGACCCCGATCACGA TGGAAGAGCTGGAGGAGATGCTGATGGACTTCGGCATCATGAAGTCGGA GGAGCAGCAGCTCGCCGAGCTCCAGGCCAAGGAAGCCGCCAAGGCCTGA	Rhodospirillales (Alphaproteobacteria)
Azospirillum canadense LMG 23617	GU256446	Rhodospirillales (Alphaproteobacteria)
Azospirillum doebereinerae DSM 13131	FJ799358	Rhodospirillales (Alphaproteobacteria)
Azospirillum halopraeferens DSM 3675	GU256447	Rhodospirillales (Alphaproteobacteria)
Azospirillum irakense DSM 11586	GU256448	Rhodospirillales (Alphaproteobacteria)
Azospirillum lipoferum DSM 1691	GU256449	Rhodospirillales (Alphaproteobacteria)
Azospirillum melinis LMG 24250	GU256450	Rhodospirillales (Alphaproteobacteria)

Azospirillum oryzae JCM 21588	AB594477	Rhodospirillales (Alphaproteobacteria)
Azospirillum picis DSM 19922	GU256451	Rhodospirillales (Alphaproteobacteria)
Azospirillum rugosum DSM 19657	GU256452	Rhodospirillales (Alphaproteobacteria)
Azospirillum sp. B510	AP010946	Rhodospirillales (Alphaproteobacteria)
Azospirillum zeae Gr61	FR669147	Rhodospirillales (Alphaproteobacteria)
Azospirillum zeae N7	DQ682472	Rhodospirillales (Alphaproteobacteria)
Bradyrhizobium sp. BTAi1	CP000494	Rhizobiales (Alphaproteobacteria)
Cyanothece sp. ATCC 51142	AF003336	Cyanobacteria
Magnetospirillum gryphiswaldense MSR-1	DQ482050	Rhodospirillales (Alphaproteobacteria)
Mesorhizobium loti MAFF303099	BA000012	Rhizobiales (Alphaproteobacteria)
Nostoc azollae' 0708	CP002059	Cyanobacteria
Rhizobium etli CFN 42	U80928	Rhizobiales (Alphaproteobacteria)
Rhodopseudomonas palustris CGA009	BX572607	Rhizobiales (Alphaproteobacteria)
Rhodospirillum centenum SW	CP000613	Rhodospirillales (Alphaproteobacteria)
Rhodospirillum rubrum ATCC 11170	CP000230	Rhodospirillales (Alphaproteobacteria)
Sinorhizobium meliloti 1021	AE006469	Rhizobiales (Alphaproteobacteria)
Xanthobacter autotrophicus Py2	CP000781	Rhizobiales (Alphaproteobacteria)

DISCUSSÃO

Apesar de *A. amazonense* ter um potencial biotecnológico considerável, o escasso conhecimento sobre as bases moleculares de sua biologia pode dificultar a identificação e o entendimento dos mecanismos responsáveis pelo efeito sobre o crescimento vegetal. Dois fatores principais até pouco tempo intrincavam o estudo da biologia molecular desse organismo: a ausência de genoma sequenciado e de ferramentas genéticas específicas. Portanto, neste estudo, esforços visando mitigar esses problemas foram realizados.

Previamente ao sequenciamento do genoma de *A. amazonense*, poucas sequências dessa bactéria estavam disponibilizadas publicamente no GenBank, sendo que a maioria delas era de rDNA 16S. Devido a essa limitação no estudo de *A. amazonense*, em princípio nosso grupo de pesquisa se deteve no isolamento de genes de interesse dessa espécie. Os primeiros genes isolados dessa espécie foram genes que codificavam para as proteínas GlnK e GlnB (SANT'ANNA, 2007; TRENTINI, 2007). Ambas proteínas são conhecidas como proteínas PII e têm papel-chave na regulação do metabolismo do nitrogênio (ARCONDÉGUY *et al.*, 2001; LEIGH & DODSWORTH, 2007; SANT'ANNA *et al.*, 2009). Entretanto, as funções dessas proteínas não foram avaliadas em *A. amazonense*, tendo em vista que não havia recursos técnicos para sua caracterização.

Embora as sequências gênicas já revelem informações importantes sobre a biologia dos organismos, a experimentação *in vivo* é uma necessidade para se correlacionar funções específicas a determinadas regiões de DNA. Consequentemente, um dos objetivos principais do presente trabalho foi otimizar técnicas moleculares para a manipulação genética de *A. amazonense*.

Toda manipulação genética depende da transferência de DNA recombinante para o organismo de interesse. Em virtude disso, metodologias de transferência de DNA para *A. amazonense* foram testadas. Os vetores pPZPLACEYFP e pHRGFPGUS foram utilizados, pois possuem as seguintes características em comum:

- marca de resistência à canamicina, que possibilitaria a seleção das bactérias que receberam os vetores;

- genes que codificam proteínas fluorescentes, propriedade adicional que permitiria uma rápida avaliação da transferência de DNA. Cada vetor codifica uma proteína distinta, pPZPLACEYFP produz a proteína EYFP e pHRGFPGUS produz a proteína EGFP;

- sítios de mobilização, ou seja, que possibilitariam a transferência por conjugação mediada por *E. coli*;

- origens de replicação de amplo espectro. O vetor pPZPLACEYFP possui a origem de replicação pVS1 e o vetor pHRGFPGUS possui a origem de replicação pBBR1.

Os plasmídeos foram transferidos para *A. amazonense* através de conjugação triparental mediada por *E. coli*. Os transconjugantes foram confirmados por apresentarem resistência à canamicina e fluorescência, demonstrando que os vetores se replicam apropriadamente nas células receptoras.

Como a metodologia de conjugação demanda um tempo considerável, também foi testada a possibilidade de se transferir DNA para *A. amazonense* através de eletroporação. Inicialmente, empregou-se uma metodologia originalmente descrita para *A. brasilense* (VAN DE BROEK *et al.*, 1989). As células de *A. amazonense* cultivadas em diferentes fases de crescimento foram lavadas e ressuspendidas em glicerol 10% e posteriormente transformadas utilizando diferentes parâmetros elétricos. Entretanto, apesar de vários parâmetros terem sido testados, os resultados decorrentes foram negativos. Van de Broek *et al.* (1989) relataram que a eficiência de transformação entre as linhagens de *Azospirillum* era variável e que esse fato poderia ser uma consequência das diferentes constituições de membrana.

Como alternativa foi também testado um protocolo de eletrotransformação de *Magnetospirillum gryphiswaldense* (bactéria da mesma família de *A. amazonense*) (SCHULTHEISS & SCHULER, 2003), que utiliza uma composição de tampão contendo sacarose e $MgCl_2$, compostos que atuariam como agentes estabilizadores de membrana (WANG & GRIFFITHS, 2009). Através dessa metodologia de eletroporação foi possível obter transformantes de *A. amazonense*. Diferentes parâmetros de eletroporação foram testados e os que tiveram melhores resultados - fase log inicial de cultivo, 200 ohms, 12,5 kV/cm - não diferiram dos geralmente utilizados em outras bactérias (LURQUIN, 1997; AUNE & AACHMANN, 2010). Considerando que espécies relacionadas à *A. amazonense* são recalcitrantes a metodologias de transformação (SCHULTHEISS & SCHULER, 2003; VAN DE BROEK *et al.*, 1989), pode-se concluir que o protocolo de eletroporação otimizado para *A. amazonense* foi um sucesso relativo. Embora as eficiências não tenham sido elevadas, foram suficientes para que os objetivos subsequentes fossem concretizados.

Uma vez que metodologias de transferência de DNA foram estabelecidas, visou-se desenvolver um protocolo de mutagênese sítio-dirigida para *A. amazonense*. A metodologia empregada também foi baseada em uma técnica usualmente utilizada em *A. brasilense*. O gene alvo *glnK* foi interrompido por um cassete de resistência à canamicina e ligado em um vetor derivado do plamídeo pSUP202 que continha uma marca de resistência a gentamicina (dados não mostrados). A ideia original seria transferir essa construção para *A. amazonense* e isolar linhagens que fossem sensíveis à gentamicina e resistentes à canamicina, oriundas da substituição do alelo selvagem pelo mutante. A integração do plasmídeo ao genoma da bactéria foi obtida facilmente (dados não mostrados), entretanto o segundo evento de recombinação, que permitiria a substituição do gene alvo e a excisão do plasmídeo, não foi detectado, apesar de milhares de colônias terem sido analisadas.

Alternativamente, empregou-se outra metodologia de mutagênese sítio-dirigida baseada em um protocolo desenvolvido para *Corynebacterium glutamicum* (EGGELING & REYES, 2005). Primeiramente, a técnica de *Crossover PCR* foi empregada para gerar uma deleção em fase no interior do gene *glnK*. A vantagem dessa metodologia em relação ao usualmente empregado em *A. brasilense*, que se vale da inserção de cassetes de resistência nos genes-alvo (LINK *et al.*, 1997), é que teoricamente os mutantes gerados não apresentariam efeitos polares decorrentes da disrupção gênica. Subsequentemente, o produto de amplificação contendo a deleção do gene *glnK* foi ligado ao vetor pK19MOBSACB. Esse plasmídeo possui a marca de contra-seleção *sacB*, que codifica a enzima levana sacarase que é letal na presença de sacarose. Sugere-se que a levana produzida a partir da sacarose se acumule no periplasma, o que seria tóxico às bactérias gram-negativas (REYRAT *et al.*, 1998). A função primordial dessa marca é facilitar a contra-seleção de linhagens que realizaram o segundo evento de recombinação, ou seja, as bactérias que se desenvolverem na presença de sacarose provavelmente substituíram o alelo selvagem e tiveram o plasmídeo removido do genoma.

A construção resultante, denominada pKΔK, foi transferida para a bactéria por conjugação e uma colônia resistente à canamicina e sensível à sacarose foi isolada, confirmando que o vetor foi integrado ao genoma. Em seguida, essa a linhagem merodiplóide, ou seja, que possui ambos alelos, selvagem e mutante, foi cultivada por várias gerações objetivando a eliminação do vetor integrado e a substituição do alelo selvagem pelo mutante. Como consequência, colônias resistentes à sacarose e sensíveis à

canamicina foram isoladas, indicando que mutantes em fase de *A. amazonense* foram gerados. A análise dos mutantes por PCR confirmou que esse sistema de mutagênese sítio-dirigida pode ser satisfatoriamente empregado em *A. amazonense*.

Neste trabalho, também foi desenvolvido um sistema-repórter baseado na expressão da proteína fluorescente EYFP. Com esse sistema aliado à análise *in silico*, foi possível detectar promotores sigma 70 putativos de *A. amazonense*. Sua aplicação principal seria a investigação de elementos envolvidos na regulação transcricional.

A otimização de técnicas de manipulação genética para *A. amazonense* viabiliza inúmeras possibilidades. Por exemplo, em *A. brasilense* já foram realizados experimentos de expressão heteróloga de genes de interesse. Em 2004, Jayaraj *et al.* (2004) transferiram para essa espécie um plasmídeo que codificava uma quitinase vegetal. Como resultado, os transconjugantes apresentaram atividade inibitória contra fungos patógenos de plantas. Em outro estudo, um plasmídeo contendo o gene da enzima 1-aminociclopropano-1-carboxilato (ACC) desaminase de *Enterobacter cloacae* foi transferido para *A. brasilense*. Essa enzima tem a propriedade de degradar o precursor de etileno, fitormônio que tem efeito inibitório na expansão celular vegetal. As bactérias que receberam o vetor foram capazes de promover o crescimento de plântulas de tomate (HOLGUIN & GLICK, 2001). Além desses, um trabalho recente demonstrou que plantas de trigo tiveram sua biomassa aumentada ao serem inoculadas com linhagens de *A. brasilense* contendo vetores que expressavam o gene *ipdC* (envolvido na produção do fitormônio estimulador ácido indol-acético, AIA) (BAUDOIN *et al.*, 2010). Até o momento, trabalhos desse tipo ainda são ocasionais, mas demonstram o potencial que a engenharia genética oferece. Embora, em geral, exista considerável resistência ao uso de organismos geneticamente modificados, não seria razoável descartar a perspectiva de se desenvolver linhagens de *A. amazonense* mais eficientes na estimulação do crescimento vegetal.

A segunda parte do trabalho se deteve na análise do genoma de *A. amazonense*. A publicação do *draft* do genoma dessa bactéria foi possível graças ao advento de tecnologias de sequenciamento mais eficientes, no caso, o pirosequenciamento. Entretanto, em virtude dos problemas intrínsecos da técnica (KIRCHER & KELSO, 2010) e da complexidade do genoma em questão, que é relativamente grande e rico em sequências repetitivas, a montagem das sequências está longe do ideal. Um fato que ilustra bem esse problema é o número elevado de *contigs* que compõem esse genoma. Além disso, o tamanho do genoma

predito de 7,3 Mpb (MARTIN-DIDONET *et al.*, 2000) não condiz com o tamanho do genoma sequenciado de 7 Mpb. A principal causa disso é que muitos *reads* de sequências repetitivas, como de transposases e de rDNA ribossômico, ficaram condensados em *contigs* únicos no processo de montagem (Luiz Gonzaga Paula de Almeida, LNCC, comunicação pessoal). Portanto, no estágio em que se encontra o genoma, análises quantitativas dos elementos repetitivos ainda não são confiáveis. No entanto, a cobertura de sequenciamento estimada em 35 x garante a acurácia das sequencias disponibilizadas.

Por enquanto, foram anotadas cerca de 3300 CDSs no genoma de *A. amazonense*. Esse número está ainda muito abaixo do número de CDSs presentes em genomas de tamanho semelhante ao de *A. amazonense*, como por exemplo, os genomas de *A. brasilense* Sp245 com 6927 CDSs distribuídas em 7,5 Mpb (http://genome.ornl.gov/microbial/abra/19sep08/) e de *Azospirillum* sp. B510 com 6309 CDSs em 7,6 Mpb (KANEKO *et al.*, 2010). Portanto, esforços ainda têm de ser realizados para melhorar a qualidade da montagem e da anotação do genoma de *A. amazonense*.

Mesmo com esses problemas, a quantidade de informação adquirida pelo sequenciamento do genoma de *A. amazonense* foi significativa. Através de uma abordagem direcionada a genes de interesse, foi possível elucidar aspectos relevantes de sua evolução e fisiologia.

Do ponto de vista evolutivo, seria importante destacar dois pontos capitais. Primeiro, a maior semelhança de grande parte dos genes de *A. amazonense* com os de *R. centenum*. Essa observação confirma resultados de análises filogenéticas usando rDNA 16S, que demonstraram a estreita relação entre ambas espécies bacterianas (FANI *et al.*, 1995; ZHANG *et al.*, 2003). A maior implicação desses resultados é que futuramente a taxonomia de *A. amazonense* e *R. centenum* tenha de ser revista.

O segundo ponto importante está relacionado com alguns genes de *A. amazonense* que são mais similares com os de membros da ordem Rhizobiales (famílias Bradyrhizobiaceae e Xanthobacteriaceae) do que com os de integrantes da própria ordem. Um exemplo notável corresponde ao sistema de fixação de nitrogênio de *A. amazonense*. Em relação a esse caso particular, é tentadora a hipótese de que tenha ocorrido um evento de transferência horizontal gênica entre o ancestral de *A. amazonense* e um membro da ordem Rhizobiales. Essa possibilidade não seria um caso isolado, já que diferentes estudos já demonstraram a transferência horizontal de genes da nitrogenase entre bactérias

(BOLHUIS *et al.*, 2009; CANTERA *et al.*, 2004). De fato, em procariotos, quando existe uma incongruência entre a história filogenética da espécie em relação a do gene, a transferência lateral de genes é a primeira conjectura considerada (DAUBIN *et al.*, 2003).

Se confirmada essa possibilidade, mais uma vez ficaria exemplificada a plasticidade dos genomas bacterianos, que são receptivos a inovações genômicas em razão da dinâmica do ambiente em que vivem (GOGARTEN & TOWNSEND, 2005). Entretanto, mais evidências seriam necessárias para descartar um provável caso de pseudo-ortologia (GOGARTEN & TOWNSEND, 2005), ou seja, onde uma duplicação gênica ocorrida no ancestral seguida de perdas gênicas diferenciais em cada linhagem mais recente também poderia explicar a topologia da árvore de *nifH*. Talvez a consequência maior da identificação dessas incongruências evolutivas é que a espécie *A. brasilense* nem sempre seria o modelo mais adequado para se basear.

Do ponto de vista fisiológico, o genoma contribuiu para se identificar características do metabolismo de *A. amazonense* antes desconhecidas. A descoberta mais importante foi a localização de genes que codificam para a enzima RubisCO, envolvida na fixação de carbono. Portanto, a bactéria em questão poderia ter hábito autotrófico, já que utilizaria gás carbônico como fonte de carbono para sua subsistência.

Outras revelações dignas de menção foram as contradições encontradas entre os dados gerados pelo sequenciamento em relação aos de estudos prévios. Por exemplo, o estudo de Martínez-Drets *et al.* (1985) sugeriu que *A. amazonense* não possuiria a via glicolítica no metobolismo do carbono, visto que não foram detectadas as atividades das enzimas fosfofrutoquinase e frutose bisfosfato aldolase em seu extrato celular. Entretanto os genes que codificam para tais proteínas estão presentes no genoma. Similarmente, apesar de o genoma apresentar um gene putativo que codificaria para uma provável homoserinalactona sintetase, o estudo de Vial *et al*. (2006) não detectou homoserina lactonas na cultura de *A. amazonense* Y2. Portanto, ao menos duas possibilidades explicariam esses resultados discrepantes: a ocorrência de artefatos técnicos que comprometeram as atividades das enzimas ou que impediram a sua detecção, ou talvez a existência de algum problema particular da espécie que impeça a apropriada expressão ou atividade dessas enzimas, como por exemplo, mutações em sítios críticos. Além desses casos, chamou atenção a presença de três genes que codificam proteínas PII no genoma, apesar de anteriormente se considerar que havia apenas dois deles (POTRICH *et al.*, 2001;

SANT'ANNA, 2007). Essa seria uma característica diferencial em relação à espécie *A. brasilense*, possuidora de duas proteínas PII, que denotaria uma maior complexidade do sistema de regulação do metabolismo nitrogênio, visto que essas proteínas interagem com diversas proteínas-alvo.

Além dos aspectos evolutivos e fisiológicos, o genoma também revelou genes que podem ter influência direta e indireta sobre a capacidade de *A. amazonense* em promover o crescimento vegetal. Concernente aos mecanismos diretos de promoção, além dos genes de fixação, foi identificado um gene que codifica para uma provável nitrilase, que poderia estar envolvida na síntese de AIA, composto conhecido por estar envolvido na regulação do crescimento vegetal (KRIECHBAUMER *et al.*, 2007). Um recente estudo realizado com *A. brasilense* demonstrou que esse fitormônio não atua somente na planta, mas também altera a expressão de vários genes bacterianos, como por exemplo, os que codificam o sistema de secreção do tipo VI, envolvido na interação de bactérias com eucariotos (VAN PUYVELDE *et al.*, 2011). Seria relevante verificar futuramente se esse gene de *A. amazonense* está realmente envolvido na síntese de ácido indolacético e caso confirmado, avaliar suas influências diretas.

Ainda que mecanismos indiretos de promoção do crescimento vegetal não tenham sido avaliados em *A. amazonense*, foram detectados genes que poderiam atuar no antagonismo de organismos patogênicos. Nesse sentido destaca-se o gene que codifica para uma provável homoserinalactonase. Tendo em vista que as bactérias patogênicas dependem de *quorum sensing* para exercer sua virulência, essa enzima impediria que esse processo ocorresse pela degradação de homoserinalactonas (LOH *et al.*, 2002; PARK *et al.*, 2003).

Embora a análise do genoma seja preliminar, já demonstrou a complexidade e versatilidade do genoma de *A. amazonense*. Cabe reiterar a necessidade de estudos para se compreender funcionalmente as sequências disponibilizadas. Por exemplo, das sequências codificantes anotadas, pelo menos um terço tem função desconhecida. Além disso, uma vez que as funções dos genes, durante o processo de anotação do genoma, são designadas através de análises comparativas *in silico*, as ferramentas genéticas otimizadas nesse trabalho serão aliadas fundamentais para se confirmar as predições e para se descobrir novas funções gênicas. Nesse sentido, a metodologia de mutagênese sítio-dirigida será

indispensável. Além disso, o sistema repórter desenvolvido poderá ser utilizado na investigação de promotores e sítios reguladores preditos no genoma de *A. amazonense*.

Neste trabalho, avanços significativos concernentes ao estudo da biologia da espécie *A. amazonense* foram estabelecidos. Com isso, espera-se que futuramente as bases moleculares do processo de promoção do crescimento vegetal sejam melhor compreendidas. Contudo, tendo em vista a alta complexidade da biologia molecular desse microrganismo, estudos têm que ser intensificados para que isso se concretize.

PERSPECTIVAS

- Otimizar a montagem do genoma, visando diminuir ao máximo o número de *contigs*;

- Analisar continuamente a anotação dos genes;

- Reclassificar taxonomicamente os integrantes do clado que contém as bactérias *A. amazonense*, *A. irakense*, *Rhodocista* spp. e *Skermanella* spp.;

- Expressar e caracterizar a enzima putativa nitrilase;

- Expressar e caracterizar a enzima putativa homoserinalactonase;

- Construir e avaliar mutantes de genes potencialmente envolvidos no processo de estimulação do crescimento vegetal, principalmente os da fixação do nitrogênio e da síntese de fitormônios;

- Identificar *regulons* de interesse, ou seja, identificar elementos transcricionais trans e cis-atuantes e os genes por eles regulados;

- Identificar genes expressos durante o processo de fixação do nitrogênio através de análises proteômicas e transcriptômicas;

- Avaliar se *A. amazonense* é capaz de antagonizar patógenos vegetais;

- Avaliar a capacidade de *A. amazonense* em fixar gás carbônico;

- Reavaliar a capacidade de *A. amazonense* em utilizar glicólise através da via glicolítica;

- Reavaliar a capacidade de *A. amazonense* em realizar *quorum sensing*;

- Avaliar a importância da transferência horizontal gênica no processo evolutivo de *A. amazonense*.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Antoun, H. & Prévost, D. Ecology of plant growth promoting rhizobacteria In: Siddiqui, Z. A., editor. PGPR: Biocontrol and Biofertilization. Netherlands: Springer, 2005. p. 1-38.

2. ARAÚJO, L. M.; HUERGO, L. F.; INVITTI, A. L.; GIMENES, C. I.; BONATTO, A. C.; MONTEIRO, R. A.; SOUZA, E. M.; PEDROSA, F. O., & CHUBATSU, L. S. Different responses of the GlnB and GlnZ proteins upon in vitro uridylylation by the *Azospirillum brasilense* GlnD protein. Brazilian Journal of Medical and Biological Research, 41(4): 289-294, 2008.

3. ARAÚJO, L. M.; MONTEIRO, R. A.; SOUZA, E. M.; STEFFENS, M. B. R.; RIGO, L. U.; PEDROSA, F. O., & CHUBATSU, L. S. GlnB is specifically required for *Azospirillum brasilense* NifA activity in *Escherichia coli*. Research in Microbiology, 155(6): 491-495, 2004.

4. ARCONDÉGUY, T.; JACK, R., & MERRICK, M. PII signal transduction proteins, pivotal players in microbial nitrogen control. Microbiology and Molecular Biology Reviews, 65(1): 80-105, 2001.

5. ARSENE, F.; KAMINSKI, P. A., & ELMERICH, C. Modulation of NifA activity by PII in *Azospirillum brasilense*: evidence for a regulatory role of the NifA N-terminal domain. Journal of Bacteriology, 178(16): 4830-4838, 1996.

6. AUNE, T. E. V. & AACHMANN, F. L. Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. Applied Microbiology and Biotechnology, 85(5): 1301-1313, 2010.

7. BALDANI, J. I. & BALDANI, V. History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience. Anais da Academia Brasileira de Ciências, 77(3): 549-579, 2005.

8. BARNEY, B. M.; YURTH, M. G.; DOS SANTOS, P. C.; DEAN, D. R., & SEEFELDT, L. C. A substrate channel in the nitrogenase MoFe protein. Journal of Biological Inorganic Chemistry, 14(7): 1015-1022, 2009.

9. BASHAN, Y. & DE-BASHAN, L. E. How the plant growth-promoting bacterium *Azospirillum* promotes plant growth: a critical assessment. Advances in Agronomy, 10877-136, 2010.

10. BAUDOIN, E.; LERNER, A.; MIRZA, S.; EL ZEMRANY, H.; PRIGENT-COMBARET, C.; JURKEVICH, E.; SPAEPEN, S.; VANDERLEYDEN, J.; NAZARET, S.; OKON, Y., & MOËNNE-LOCCOZ, Y. Effects of *Azospirillum brasilense* with genetically modified auxin biosynthesis gene *ipdC* upon the diversity of the indigenous microbiota of the wheat rhizosphere. Research in Microbiology, 161(3): 219-226, 2010.

11. BERG, G. Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. Applied Microbiology and Biotechnology, 84(1): 11-18, 2009.

12. BOLHUIS, H.; SEVERIN, I.; CONFURIUS-GUNS, V.; WOLLENZIEN, U. I. A., & STAL, L. J. Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium *Microcoleus chthonoplastes*. The ISME Journal, 4(1): 121-130, 2009.

13. CANTERA, J. J.; KAWASAKI, H., & SEKI, T. The nitrogen-fixing gene (*nifH*) of *Rhodopseudomonas palustris*: a case of lateral gene transfer? Microbiology (Reading, England), 150(Pt 7): 2237-2246, 2004.

14. CHEN, S.; LIU, L.; ZHOU, X.; ELMERICH, C., & LI, J. L. Functional analysis of the GAF domain of NifA in *Azospirillum brasilense*: effects of Tyr-->Phe mutations on NifA and its interaction with GlnB. Molecular Genetics and Genomics: MGG, 273(5): 415-422, 2005.

15. DAUBIN, V.; MORAN, N. A., & OCHMAN, H. Phylogenetics and the cohesion of bacterial genomes. Science, 301(5634): 829-832, 2003.

16. DAVISON, J. Genetic tools for pseudomonads, rhizobia, and other gram-negative bacteria. BioTechniques, 32(2): 2002.

17. DE ZAMAROCZY, M.; PAQUELIN, A.; PELTRE, G.; FORCHHAMMER, K., & ELMERICH, C. Coexistence of two structurally similar but functionally different PII proteins in *Azospirillum brasilense*. Journal of Bacteriology, 178(14): 4143-4149, 1996.

18. DE ZAMAROCZY, M. Structural homologues P(II) and P(Z) of *Azospirillum brasilense* provide intracellular signalling for selective regulation of various nitrogen-dependent functions. Molecular Microbiology, 29(2): 449-463, 1998.

19. DEKHIL, S. B.; CAHILL, M.; STACKEBRANDT, E., & SLY, L. I. Transfer of *Conglomeromonas largomobilis* subsp *largomobilis* to the genus *Azospirillum* as *Azopirillum largomobile* comb nov, and elevation of *Conglomeromonas largomobilis* subsp *parooensis* to the new type species of *Conglomeromonas*, *Conglomeromonas parooensi*. Systematic and Applied Microbiology, 2072-77, 1997.

20. DIXON, R. & KAHN, D. Genetic regulation of biological nitrogen fixation. Nature Reviews Microbiology, 2(8): 621-631, 2004.

21. ECKERT, B.; WEBER, O. B.; KIRCHHOF, G.; HALBRITTER, A.; STOFFELS, M., & HARTMANN, A. *Azospirillum doebereinera*e sp. nov., a nitrogen-fixing bacterium associated with the C4-grass *Miscanthus*. International Journal of Systematic and Evolutionary Microbiology, 51(Pt 1): 17-26, 2001.

22. EDGREN, T. & NORDLUND, S. The *fixABCX* genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase. Journal of Bacteriology, 186(7): 2052-2060, 2004.

23. Eggeling, L. & Reyes, O. Deletion of chromosomal sequences and allelic exchange In: Eggeling, L. & Bott, M., editors. Handbook of *Corynebacterium glutamicum*. Boca Raton, FL: CRC press, 2005. p. 557-559.

24. FANI, R.; BANDI, C.; BAZZICALUPO, M.; CECCHERINI, M. T.; FANCELLI, S.; GALLORI, E.; GERACE, L.; GRIFONI, A.; MICLAUS, N., & DAMIANI, G. Phylogeny of the genus *Azospirillum* based on 16S rDNA sequence. FEMS Microbiology Letters, 129(2-3): 195-200, 1995.

25. FISCHER, H. M.; BRUDERER, T., & HENNECKE, H. Essential and non-essential domains in the *Bradyrhizobium japonicum* NifA protein: identification of indispensable cysteine residues potentially involved in redox reactivity and/or metal binding. Nucleic Acids Research, 16(5): 2207-2224, 1988.

26. FRAZZON, J. & SCHRANK, I. S. Sequencing and complementation analysis of the nifUSV genes from Azospirillum brasilense. FEMS Microbiology Letters, 159(2): 151-158, 1998.

27. FU, H. A.; HARTMANN, A.; LOWERY, R. G.; FITZMAURICE, W. P.; ROBERTS, G. P., & BURRIS, R. H. Posttranslational regulatory system for nitrogenase activity in *Azospirillum* spp. Journal of Bacteriology, 171(9): 4679-4685, 1989.

28. Fuentes-Ramirez, L. E. & Caballero-Mellado, J. Bacterial biofertilizers In: Siddiqui, Z. A., editor. PGPR: Biocontrol and Biofertilization. Dordrecht, The Netherlands: Springer, 2005. p. 143-172.

29. GOGARTEN, J. P. & TOWNSEND, J. P. Horizontal gene transfer, genome innovation and evolution. Nature Reviews Microbiology, 3(9): 679-687, 2005.

30. HARTMANN, A.; FU, H., & BURRIS, R. H. Regulation of nitrogenase activity by ammonium chloride in *Azospirillum* spp. Journal of Bacteriology, 165(3): 864-870, 1986.

31. HOLGUIN, G. & GLICK, B. R. Expression of the ACC Deaminase gene from *Enterobacter cloacae* UW4 in *Azospirillum brasilense*. Microbial Ecology, 41(3): 281-288, 2001.

32. HOLGUIN, G.; PATTEN, C. L., & GLICK, B. R. Genetics and molecular biology of *Azospirillum*. Biology and Fertility of Soils, 29(1): 10-23, 1999.

33. HUANG, J.; PRAY, C., & ROZELLE, S. Enhancing the crops to feed the poor. Nature, 418(6898): 678-684, 2002.

34. HUERGO, L. F.; MERRICK, M.; MONTEIRO, R. A.; CHUBATSU, L. S.; STEFFENS, M. B.; PEDROSA, F. O., & SOUZA, E. M. In vitro interactions between the PII proteins and the nitrogenase regulatory enzymes dinitrogenase reductase ADP-ribosyltransferase (DraT) and dinitrogenase reductase-activating glycohydrolase (DraG) in *Azospirillum brasilense*. The Journal of Biological Chemistry, 284(11): 6674-6682, 2009.

35. HUERGO, L. F.; MERRICK, M.; PEDROSA, F. O.; CHUBATSU, L. S.; ARAUJO, L. M., & SOUZA, E. M. Ternary complex formation between AmtB, GlnZ and the nitrogenase regulatory enzyme DraG reveals a novel facet of nitrogen regulation in bacteria. Molecular Microbiology, 66(6): 1523-1535, 2007.

36. HUERGO, L. F.; SOUZA, E. M.; ARAUJO, M. S.; PEDROSA, F. O.; CHUBATSU, L. S.; STEFFENS, M. B., & MERRICK, M. ADP-ribosylation of dinitrogenase reductase in *Azospirillum brasilense* is regulated by AmtB-dependent membrane sequestration of DraG. Molecular Microbiology, 59(1): 326-337, 2006.

37. HUERGO, L. F.; SOUZA, E. M.; STEFFENS, M. B. R.; YATES, G.; PEDROSA, F. O., & CHUBATSU, L. S. Regulation of *glnB* gene promoter expression in *Azospirillum brasilense* by the NtrC protein. FEMS Microbiology Letters, 223(1): 33-40, 2003.

38. KANEKO, T.; MINAMISAWA, K.; ISAWA, T.; NAKATSUKASA, H.; MITSUI, H.; KAWAHARADA, Y.; NAKAMURA, Y.; WATANABE, A.; KAWASHIMA, K.; ONO, A.; SHIMIZU, Y.; TAKAHASHI, C.; MINAMI, C.; FUJISHIRO, T.; KOHARA, M.; KATOH, M.; NAKAZAKI, N.; NAKAYAMA, S.; YAMADA, M.; TABATA, S., & SATO, S. Complete genomic structure of the cultivated rice endophyte *Azospirillum* sp. B510. DNA Research, 17(1): 37-50, 2010.

39. KHAMMAS, K. M.; AGERON, E.; GRIMONT, P. A., & KAISER, P. *Azospirillum irakense* sp. nov., a nitrogen-fixing bacterium associated with rice roots and rhizosphere soil. Research in Microbiology, 140(9): 679-693, 1989.

40. KIRCHER, M. & KELSO, J. High-throughput DNA sequencing--concepts and limitations. BioEssays, 32(6): 524-536, 2010.

41. KLOEPPER, J. W.; LIFSHITZ, R., & ZABLOTOWICZ, R. M. Free-living bacterial inocula for enhancing crop productivity. Trends in Biotechnology, 7(2): 39-44, 1989.

42. KRIECHBAUMER, V.; PARK, W. J.; PIOTROWSKI, M.; MEELEY, R. B.; GIERL, A., & GLAWISCHNIG, E. Maize nitrilases have a dual role in auxin homeostasis and beta-cyanoalanine hydrolysis. Journal of Experimental Botany, 58(15-16): 4225-4233, 2007.

43. LAVRINENKO, K.; CHERNOUSOVA, E.; GRIDNEVA, E.; DUBININA, G.; AKIMOV, V.; KUEVER, J.; LYSENKO, A., & GRABOVICH, M. *Azospirillum thiophilum* sp. nov., a diazotrophic bacterium isolated from a sulfide spring. International Journal of Systematic and Evolutionary Microbiology, 60(Pt 12): 2832-2837, 2010.

44. LEIGH, J. A. & DODSWORTH, J. A. Nitrogen regulation in bacteria and archaea. Annual Review of Microbiology, 61349-377, 2007.

45. LERNER, A.; CASTRO-SOWINSKI, S.; LERNER, H.; OKON, Y., & BURDMAN, S. Glycogen phosphorylase is involved in stress endurance and biofilm formation in *Azospirillum brasilense* Sp7. FEMS Microbiology Letters, 300(1): 75-82, 2009.

46. LERNER, A.; CASTRO-SOWINSKI, S.; VALVERDE, A.; LERNER, H.; DROR, R.; OKON, Y., & BURDMAN, S. The *Azospirillum brasilense* Sp7 *noeJ* and *noeL* genes are involved in extracellular polysaccharide biosynthesis. Microbiology (Reading, England), 155(Pt 12): 4058-4068, 2009.

47. LI, H.; XU, F.; REN, X., & CHEN, S. Functional analysis of the *fixL/fixJ* and *fixK* genes in *Azospirillum brasilense* Sp7. Annals of Microbiology, 60(3): 469-480, 2010.

48. LIN, S. Y.; YOUNG, C. C.; HUPFER, H.; SIERING, C.; ARUN, A. B.; CHEN, W. M.; LAI, W. A.; SHEN, F. T.; REKHA, P. D., & YASSIN, A. F. *Azospirillum picis* sp. nov., isolated from discarded tar. International Journal of Systematic and Evolutionary Microbiology, 59(Pt 4): 761-765, 2009.

49. LIN, S. Y.; SHEN, F. T.; YOUNG, L. S.; ZHU, Z. L.; CHEN, W. M., & YOUNG, C. C. *Azospirillum formosense* sp. nov., a novel diazotrophic bacterium isolated from agricultural soil. International Journal of Systematic and Evolutionary Microbiology, 2011.

50. LINK, A. J.; PHILLIPS, D., & CHURCH, G. M. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. Journal of Bacteriology, 179(20): 6228-6237, 1997.

51. LOH, J.; PIERSON, E. A.; PIERSON, L. S. I.; STACEY, G., & CHATTERJEE, A. Quorum sensing in plant-associated bacteria. Current Opinion in Plant Biology, 5(4): 285-290, 2002.

52. LU, Y. K.; MARDEN, J.; HAN, M.; SWINGLEY, W. D.; MASTRIAN, S. D.; CHOWDHURY, S. R.; HAO, J.; HELMY, T.; KIM, S.; KURDOGLU, A. A.; MATTHIES, H. J.; ROLLO, D.; STOTHARD, P.; BLANKENSHIP, R. E.; BAUER, C. E., & TOUCHMAN, J. W. Metabolic flexibility revealed in the genome of the cyst-forming alpha-1 proteobacterium *Rhodospirillum centenum*. BMC Genomics, 112010.

53. LUGTENBERG, B. & KAMILOVA, F. Plant-growth-promoting rhizobacteria. Annual Review of Microbiology, 63541-556, 2009.

54. LURQUIN, P. F. Gene transfer by electroporation. Molecular Biotechnology, 7(1): 5-35, 1997.

55. MAGALHÃES, F. M. M.; BALDANI, J. I.; SOUTO, S. M.; KUYKENDALL, J. R., & DÖBEREINER, J. A new acid-tolerant *Azospirillum* species. Anais da Academia Brasileira de Ciências, 55417-430, 1983.

56. MARTIN-DIDONET, C. C.; CHUBATSU, L. S.; SOUZA, E. M.; KLEINA, M.; REGO, F. G.; RIGO, L. U.; YATES, M. G., & PEDROSA, F. O. Genome structure of the genus *Azospirillum*. Journal of Bacteriology, 182(14): 4113-4116, 2000.

57. MARTÍNEZ-DRETS, G.; FABIANO, E., & CARDONA, A. Carbohydrate catabolism in *Azospirillum amazonense*. Applied and Environmental Microbiology, 50(1): 183-185, 1985.

58. MEHNAZ, S.; WESELOWSKI, B., & LAZAROVITS, G. *Azospirillum canadense* sp. nov., a nitrogen-fixing bacterium isolated from corn rhizosphere. International Journal of Systematic and Evolutionary Microbiology, 57(Pt 3): 620-624, 2007.

59. MEHNAZ, S.; WESELOWSKI, B., & LAZAROVITS, G. *Azospirillum zeae* sp. nov., a diazotrophic bacterium isolated from rhizosphere soil of *Zea mays*. International Journal of Systematic and Evolutionary Microbiology, 57(Pt 12): 2805-2809, 2007.

60. OLIVEIRA, M. A. S.; BAURA, V. A.; AQUINO, B.; HUERGO, L. F.; KADOWAKI, M. A. S.; CHUBATSU, L. S.; SOUZA, E. M.; DIXON, R.; PEDROSA, F. O.; WASSEM, R., & MONTEIRO, R. A. Role of conserved cysteine residues in *Herbaspirillum seropedicae* NifA activity. Research in Microbiology, 160(6): 389-395, 2009.

61. PARK, S. Y.; LEE, S.; OH, T. K.; OH, J. W.; KOO, B. T.; YUM, D. Y., & LEE, J. K. AhlD, an N-acylhomoserine lactonase in *Arthrobacter* sp., and predicted homologues in other bacteria. Microbiology, 149(6): 1541-1550, 2003.

62. PASSAGLIA, L. M.; NUNES, C. P.; ZAHA, A., & SCHRANK, I. S. The nifHDK operon in the free-living nitrogen-fixing bacteria *Azospirillum brasilense* sequentially comprises genes H, D, K, an 353 bp orf and gene Y. Brazilian Journal of Medical and Biological Research, 24(7): 649-675, 1991.

63. PENG, G.; WANG, H.; ZHANG, G.; HOU, W.; LIU, Y.; WANG, E. T., & TAN, Z. *Azospirillum melinis* sp. nov., a group of diazotrophs isolated from tropical molasses grass. International Journal of Systematic and Evolutionary Microbiology, 56(Pt 6): 1263-1271, 2006.

64. POTRICH, D. P.; BRESSEL, T. A.; SCHRANK, I. S., & PASSAGLIA, L. M. Sequencing and promoter analysis of the nifENXorf3orf5fdxAnifQ operon from *Azospirillum brasilense* Sp7. Brazilian Journal of Medical and Biological Research, 34(11): 1379-1395, 2001.

65. POTRICH, D. P.; PASSAGLIA, L. M., & SCHRANK, I. S. Partial characterization of nif genes from the bacterium *Azospirillum amazonense*. Brazilian Journal of Medical and Biological Research, 34(9): 1105-1113, 2001.

66. Raven, P. H., Evert, R. F., & Eichhorn, S. E. Nutrição vegetal e solos Biologia Vegetal: Guanabara Koogan SA, 2001. p. 698-719.

67. REINHOLD, B.; HUREK, T.; FENDRIK, I.; POT, B.; GILLIS, M.; KERSTERS, K.; THIELEMANS, S., & DE LEY, J. *Azospirillum halopraeferens* sp.

nov., a Nitrogen-Fixing Organism Associated with Roots of Kallar Grass (*Leptochloa fusca* (L.) Kunth). Bacteriology, 37(1): 43-51, 1987.

68. REIS JUNIOR, F. B.; ACHADO, C. T. T.; ACHADO, A. T., & ODEK, L. Inoculação de *Azospirillum amazonense* em dois genótipos de milho sob diferentes regimes de nitrogênio. Revista Brasileira de Ciência do Solo, 32(3): 1139-1146, 2008.

69. REYRAT, J. M.; PELICIC, V.; GICQUEL, B., & RAPPUOLI, R. Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. Infection and Immunity, 66(9): 4011-4017, 1998.

70. RODRIGUES, E. P.; RODRIGUES, L. S.; DE OLIVEIRA, A. L. M.; BALDANI, V. L. D.; TEIXEIRA, K. R. S.; URQUIAGA, S., & REIS, V. M. *Azospirillum amazonense* inoculation: effects on growth, yield and N_2 fixation of rice (*Oryza sativa* L.). Plant and Soil, 302(1): 249-261, 2008.

71. RUBIO, L. M. & LUDDEN, P. W. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annual Review of Microbiology, 62(1): 93-111, 2008.

72. Sant'anna, F. H. Identificação de genes de *Azospirillum amazonense* diferencialmente expressos em resposta à disponibilidade de nitrogênio., 2007.

73. SANT'ANNA, F. H.; TRENTINI, D. B.; DE SOUTO WEBER, S.; CECAGNO, R.; DA SILVA, S. C., & SCHRANK, I. S. The PII superfamily revised: a novel group and evolutionary insights. Journal of Molecular Evolution, 68(4): 322-336, 2009.

74. SCHRANK, I. S.; ZAHA, A.; DE ARAUJO, E. F., & SANTOS, D. S. Construction of a gene library from *Azospirillum brasilense* and characterization of a recombinant containing the nif structural genes. Brazilian Journal of Medical and Biological Research, 20(3-4): 321-330, 1987.

75. SCHULTHEISS, D. & SCHULER, D. Development of a genetic system for Magnetospirillum gryphiswaldense. Archives of Microbiology, 179(2): 89-94, 2003.

76. SCHWEIZER, H. P. Bacterial genetics: past achievements, present state of the field, and future challenges. BioTechniques, 44 Supplement(4): 633-641, 2008.

77. SPEROTTO, R. A.; GROSS, J.; VEDOY, C.; PASSAGLIA, L. M., & SCHRANK, I. S. The electron transfer flavoprotein *fixABCX* gene products from *Azospirillum brasilense* show a NifA-dependent promoter regulation. Current Microbiology, 49(4): 267-273, 2004.

78. STEENHOUDT, O. & VANDERLEYDEN, J. *Azospirillum*, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. FEMS Microbiology Reviews, 24(4): 487-506, 2000.

79. SUTTON, M. A.; OENEMA, O.; ERISMAN, J. W.; LEIP, A.; VAN GRINSVEN, H., & WINIWARTER, W. Too much of a good thing. Nature, 472159-161, 2011.

80. TAL, S. & OKON, Y. Production of the reserve material poly-ß-hydroxybutyrate and its function in *Azospirillum brasilense* Cd. Canadian Journal of Microbiology, 31(7): 608-613, 1985.

81. TARRAND, J. J.; KRIEG, N. R., & DÖBEREINER, J. A taxonomic study of the *Spirillum lipoferum* group, with descriptions of a new genus, *Azospirillum* gen. nov. and two species, *Azospirillum lipoferum* (Beijerinck) comb. nov. and *Azospirillum brasilense* sp. nov. Canadian Journal of Microbiology, 24(8): 967-980, 1978.

82. Trentini, D. B. Isolamento e análise do gene *glnB* de *Azospirillum amazonense*., 2007.

83. VALVERDE, A.; OKON, Y., & BURDMAN, S. cDNA-AFLP reveals differentially expressed genes related to cell aggregation of *Azospirillum brasilense*. FEMS Microbiology Letters, 265(2): 186-194, 2006.

84. VAN DE BROEK, A.; VAN GOOL, A., & VANDERLEYDEN, J. Electroporation of *Azospirillum brasilense* with plasmid DNA. FEMS Microbiology Letters, 61(1-2): 177-182, 1989.

85. VAN DOMMELEN, A.; KEIJERS, V.; SOMERS, E., & VANDERLEYDEN, J. Cloning and characterisation of the *Azospirillum brasilense glnD* gene and analysis of a *glnD* mutant. Molecular Genetics and Genomics: MGG, 266(5): 813-820, 2002.

86. VAN DOMMELEN, A.; SPAEPEN, S., & VANDERLEYDEN, J. Identification of the glutamine synthetase adenylyltransferase of *Azospirillum brasilense*. Research in Microbiology, 160(3): 205-212, 2009.

87. VAN PUYVELDE, S.; CLOOTS, L.; ENGELEN, K.; DAS, F.; MARCHAL, K.; VANDERLEYDEN, J., & SPAEPEN, S. Transcriptome analysis of the rhizosphere bacterium *Azospirillum brasilense* reveals an extensive auxin response. Microbial Ecology, 61(4): 723-728, 2011.

88. VERMA, J. P.; YADAV, J.; TIWARI, K. N.; LAVAKUSH, & SINGH, V. Impact of plant growth promoting rhizobacteria on crop production. International Journal of Agricultural Research, 5(11): 954-983, 2010.

89. VIAL, L.; CUNY, C.; GLUCHOFF-FIASSON, K.; COMTE, G.; OGER, P. M.; FAURE, D.; DESSAUX, Y.; BALLY, R., & WISNIEWSKI-DYE, F. N-acyl-homoserine lactone-mediated quorum-sensing in *Azospirillum*: an exception rather than a rule. FEMS Microbiology Ecology., 58(2): 155-168, 2006.

90. WANG, H. & GRIFFITHS, M. W. $Mg+^{2}$-free buffer elevates transformation efficiency of *Vibrio parahaemolyticus* by electroporation. Letters in Applied Microbiology, 48(3): 349-354, 2009.

91. WESTBY, C. A.; ENDERLIN, C. S.; STEINBERG, N. A.; JOSEPH, C. M., & MEEKS, J. C. Assimilation of $^{13}NH_4$+ by *Azospirillum brasilense* grown under nitrogen limitation and excess. Journal of Bacteriology, 169(9): 4211-4214, 1987.

92. XIE, C. H. & YOKOTA, A. *Azospirillum oryzae* sp. nov., a nitrogen-fixing bacterium isolated from the roots of the rice plant *Oryza sativa*. International Journal of Systematic and Evolutionary Microbiology, 55(Pt 4): 1435-1438, 2005.

93. XIE, Z.; ULRICH, L. E.; ZHULIN, I. B., & ALEXANDRE, G. PAS domain containing chemoreceptor couples dynamic changes in metabolism with chemotaxis. Proceedings of the National Academy of Sciences, 107(5): 2235-2240, 2010.

94. YOUNG, C. C.; HUPFER, H.; SIERING, C.; HO, M. J.; ARUN, A. B.; LAI, W. A.; REKHA, P. D.; SHEN, F. T.; HUNG, M. H.; CHEN, W. M., & YASSIN, A. F. *Azospirillum rugosum* sp. nov., isolated from oil-contaminated soil. International Journal of Systematic and Evolutionary Microbiology, 58(Pt 4): 959-963, 2008.

95. ZHANG, D.; YANG, H.; ZHANG, W.; HUANG, Z., & LIU, S. J. *Rhodocista pekingensis* sp. nov., a cyst-forming phototrophic bacterium from a municipal wastewater treatment plant. International Journal of Systematic and Evolutionary Microbiology, 53(4): 1111-1114, 2003.

96. ZHOU, Y.; WEI, W.; WANG, X.; XU, L., & LAI, R. *Azospirillum palatum* sp. nov., isolated from forest soil in Zhejiang province, China. The Journal of General and Applied Microbiology, 55(1): 1-7, 2009.

APÊNDICES

1- Publicação de artigo na revista Journal of Molecular Evolution (Fator de impacto: 2,31):

The PII superfamily revised: a novel group and evolutionary insights.

Fernando H Sant'Anna, Débora B Trentini, Shana S Weber, Ricardo Cecagno & Irene S Schrank.

J Mol Evol. 2009 Apr;68(4):322-36.

2- Submissão de artigo na revista DNA Research (Fator de impacto: 4,75)

Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution

Shana S Weber, Fernando H Sant'Anna & Irene S Schrank.

3- *Curriculum vitae*

Apêndice 1 - Publicação de artigo na revista Journal of Molecular Evolution

J Mol Evol (2009) 68:322–336
DOI 10.1007/s00239-009-9209-6

The PII Superfamily Revised: A Novel Group and Evolutionary Insights

Fernando Hayashi Sant'Anna · Débora Broch Trentini · Shana de Souto Weber · Ricardo Cecagno · Sérgio Ceroni da Silva · Irene Silveira Schrank

Received: 26 September 2008 / Accepted: 3 February 2009 / Published online: 19 March 2009

Abstract The PII proteins compose a superfamily of signal transducers with fundamental roles in the nitrogen metabolism of prokaryotic organisms. They act at different cellular targets, such as ammonia transporters, enzymes, and transcriptional factors. These proteins are small, highly conserved, and well distributed among prokaryotes. The current PII classification is based on sequence similarity and genetic linkage. Our work reviewed this classification through an extensive analysis of PII homologues deposited in GenBank. We also investigated evolutionary aspects of this ancient protein superfamily and revised its PROSITE signatures. A new group of PII proteins is described in this work. These PII homologues have a peculiar genetic context, as they are associated with metal transporters and do not contain the canonical PROSITE signatures of PII. Our analysis reveals that horizontal gene transfer could have played an important role in PII evolution. Thus, new insights into PII evolution, a new PII group, and more comprehensive PROSITE signatures are proposed.

Keywords PII · GlnB · GlnK · Nif · AmtB · Cellular signaling · Nitrogen metabolism · Horizontal gene transfer · Metal efflux pump

The authors Fernando Hayashi Sant'Anna and Débora Broch Trentini contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00239-009-9209-6) contains supplementary material, which is available to authorized users.

F. H. Sant'Anna · D. B. Trentini · S. de Souto Weber · R. Cecagno · S. C. da Silva
Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

I. S. Schrank (✉)
Departamento de Biologia Molecular e Biotecnologia - Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul (UFRGS), Av. Bento Gonçalves 9500, P. 43421, C.P. 15005, CEP 91501-970 Porto Alegre, RS, Brazil
e-mail: irene@cbiot.ufrgs.br

S. C. da Silva
Departamento de Patologia Clínica Veterinária, Faculdade de Veterinária, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

Introduction

The superfamily of PII signal transduction proteins is known to contain the most widely distributed signaling proteins in nature (Forchhammer 2004). The explosion of genome sequencing efforts has revealed that members of the PII family are ubiquitous among prokaryotes. Moreover, representatives of this family are spread among all domains of life, being found in nitrogen-fixing Archaea and in the chloroplasts of eukaryotic phototrophs (Arcondéguy et al. 2001; Forchhammer 2008).

PII proteins have been studied at length over the past few years, since it has become evident that these proteins play central roles in bacterial nitrogen metabolism. Their remarkable functional versatility permits them to participate in a wide range of regulatory pathways related to nitrogen control by direct interaction with a variety of cellular targets (Forchhammer 2008). The roles of PII proteins in cell metabolism have been extensively reviewed (Arcondéguy et al. 2001; Forchhammer 2008; Leigh and Dodsworth 2007; Ninfa and Atkinson 2000; Ninfa and Jiang 2005; Reitzer 2003). The crystal structure of several PII proteins has also been elucidated, consisting of homotrimeric subunits with a highly conserved three-dimensional structure (Conroy et al. 2007; Nichols et al.

Apêndice 2 - Submissão de artigo na revista DNA Research

Unveiling *Mycoplasma hyopneumoniae* promoters: sequence definition and genomic distribution

Shana de Souto Weber; Fernando Hayashi Sant'Anna & Irene Silveira Schrank

Summary

Several *Mycoplasma* species have had their genome completely sequenced, including four strains of the swine pathogen *Mycoplasma hyopneumoniae*. Nevertheless, little is known about the nucleotide sequences that control transcription initiation in these microorganisms. Therefore, with the objective of investigating the promoter sequences of *M. hyopneumoniae*, 23 transcriptional start sites (TSS) of distinct genes were mapped. A pattern that resembles the σ^{70} promoter -10 element was found upstream of the TSSs. However, no -35 element was distinguished. Instead, an AT-rich periodic signal could be observed. About half of the experimentally defined promoters possessed the string 5'-TRTGn-3', which was identical to the -16 element usually found in Gram-positive bacteria. The defined promoters were utilized to build position-specific scoring matrices (PSSM) in order to scan putative promoters upstream of all coding sequences (CDSs) of the *M. hyopneumoniae* genome. Two hundred and one signals were found distributed in 169 CDSs. Most of these sequences were located up to 100 nucleotides from the start codons. This study shows that the number of promoter-like sequences in the *M. hyopneumoniae* genome is more frequent than expected by chance, indicating that most of the sequences found are probably biologically functional.

Apêndice 3 - Curriculum vitae

Dados Pessoais

Nome Fernando Hayashi Sant'Anna

Formação Acadêmica/Titulação

2007 Doutorado em Programa de Pós-Graduação em Biologia Celular e Molecular.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
Título: Estudo da bactéria promotora de crescimento vegetal Azospirillum amazonense: aspectos genômicos e ferramentas genéticas específicas
Orientador: Irene Silveira Schrank
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2006 - 2007 Mestrado em Programa de Pós-Graduação em Biologia Celular e Molecular.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
Título: Identificação de genes diferencialmente expressos em resposta à disponibilidade de nitrogênio em Azospirillum amazonense, Ano de obtenção: 2007
Orientador: Irene Silveira Schrank
Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2001 - 2005 Graduação em Ciências Biológicas.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil
Título: ANÁLISE DO GENE glnK DA BACTÉRIA DIAZOTRÓFICA Azospirillum amazonense
Orientador: Irene Silveira Schrank
Bolsista do(a): Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul

---

Formação complementar

2004 - 2004 RDA em Microorganismos.
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brasil

Prêmios e títulos

2006 Menção Honrosa, 25ª Reunião de Genética de Microrganismos

Produção bibliográfica
Artigos completos publicados em periódicos

1. SANT'ANNA, F. H., ALMEIDA, L. G. P., CECAGNO, R., REOLON, L. A., SIQUEIRA, F. M., Machado, M. R. S., VASCONCELOS, A. T. R., SCHRANK, Irene Silveira
Genomic insights into the versatility of the plant growth-promoting bacterium Azospirillum amazonense. BMC Genomics. , v.12, p.409 - , 2011.

2. SANT'ANNA, F. H., ANDRADE, D. S., TRENTINI, D. B., WEBER, S. S., SCHRANK, Irene Silveira
Tools for genetic manipulation of the plant growth-promoting bacterium Azospirillum amazonense. BMC Microbiology (Online). , v.11, p.107 - , 2011.

3. SANT'ANNA, F. H., TRENTINI, D. B., WEBER, S. S., CECAGNO, R., SILVA, S. C., SCHRANK, Irene Silveira
The PII superfamily revised: a novel group and evolutionary insights.. Journal of Molecular Evolution. , v.68, p.322 - 336, 2009.

4. BROETTO, L., CECAGNO, R., SANT'ANNA, F. H., WEBER, S. S., SCHRANK, Irene Silveira
Stable transformation of Chromobacterium violaceum with a broad-host-range plasmid. Applied Microbiology and Biotechnology. , v.71, p.450 - 454, 2006.

Trabalhos publicados em anais de eventos (resumo)

1. SANT'ANNA, F. H., Machado, M. R. S., SCHRANK, Irene Silveira
The unusual history of nif genes from Azospirillum amazonense Y2 In: 12th International Symposium on Biological Nitrogen Fixation with Non-Legumes, 2010, Búzios.
12th International Symposium on Biological Nitrogen Fixation with Non-Legumes - Book of Abstracts. , 2010. p.21 -

2. TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Análise regulatória das proteínas PII no metabolismo de nitrogênio da bactéria diazotrófica Azospirillum amazonense In: 25 º Congresso de Brasileiro de Microbiologia, 2009, Porto de Galinhas.
25º Congresso Brasileiro de Microbiologia. , 2009.

3. ANDRADE, D. S., SANT'ANNA, F. H., TRENTINI, D. B., SCHRANK, Irene Silveira
Caracterização de genes que codificam para diguanilato ciclase/fosfodiesterase e para a proteína dnaK de Azospirillum amazonense In: 25 º Congresso de Brasileiro de Microbiologia, 2009, Porto de Galinhas.

25 º Congresso de Brasileiro de Microbiologia. , 2009.

4. Machado, M. R. S., TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Caracterização dos genes rpoH E relA em Azospirillum amazonense In: 25 º Congresso de Brasileiro de Microbiologia, 2009, Porto de Galinhas.
25 º Congresso de Brasileiro de Microbiologia. , 2009.

5. SANT'ANNA, F. H., ANDRADE, D. S., SCHRANK, Irene Silveira
Desenvolvimento de Ferramentas genéticas para Azospirillum amazonense In: 25 º Congresso de Brasileiro de Microbiologia, 2009, Porto de Galinhas.
25º Congresso Brasileiro de Microbiologia. , 2009.

6. SANT'ANNA, F. H., TRENTINI, D. B., WEBER, S. S., CECAGNO, R., SILVA, S. C., SCHRANK, Irene Silveira
Aspectos evolutivos da superfamília de proteínas sinalizadoras PII In: 26º Reunião de Genética de Microrganismos, 2008, Salvador.
26º Reunião de Genética de Microrganismos. , 2008.

7. TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Caracterização e expressão das proteínas de transdução de sinal PII do diazotrófico de vida livre Azospirillum amazonense In: 26a Reuninão de Genética de Microrganismos, 2008, Salvador.
26a Reuninão de Genética de Microrganismos. , 2008.

8. SANT'ANNA, F. H., ANDRADE, D. S., SCHRANK, Irene Silveira
Genes envolvidos na resposta de Azospirillum amazonense à limitação de nitrogênio In: 54ºCongresso Brasileiro de Genética, 2008, Salvador.
54ºCongresso Brasileiro de Genética. , 2008.

9. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Identification of differentially expressed genes of Azospirillum amazonense related to the nitrogen availability by cDNA-RDA technique In: XI Congreso Argentino de Microbiologia, 2007, Córdoba.
Revista Argentina de Microbiologia. , 2007. v.39. p.160 -

10. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Isolation of glnK gene of Azospirillum amazonense In: XI Congreso Argentino de Microbiologia, 2007, Córdoba.
Revista Argentina de Microbiologia. , 2007. v.39. p.160 -

11. TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Sequence isolation and transcriptional regulation of the PII protein encoding gene glnB of the diazotroph Azospirillum amazonense In: XI Congreso Argentino de Microbiologia, 2007, Córdoba.
Revista Argentina de Microbiologia. , 2007. v.39. p.48 -

12. TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Transcriptional regulation of the Azospirillum brasilense orf2-nifUSV-orf4 operon,

involved with nitrogenase biosynthesis In: XI Congreso Argentino de Microbiologia, 2007, Córdoba.
Revista Argentina de Microbiologia. , 2007. v.39. p.49 -

13. TRENTINI, D. B., SANT'ANNA, F. H., SCHRANK, Irene Silveira
Análise da região regulatória do operon nifUSV de Azospirillum brasilense In: XVIII Salão de iniciação científica, 2006, Porto Alegre.
XVIII Salão de iniciação científica. , 2006.

14. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Clonagem e caracterização do gene glnK de Azospirillum amazonense In: 25ª Reunião de Genética de Microrganismos, 2006, São Pedro.
25ª Reunião de Genética de Microrganismos. , 2006.

15. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Análise dos genes glnB e glnD da bactéria diazotrófica Azospirillum amazonense In: XVII Salão de Iniciação Científica, 2005, Porto Alegre.
XVII Salão de Iniciação Científica. , 2005.

16. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Análise dos genes glnB e glnD da bactéria diazotrófica Azospirillum amazonense In: 51° Congresso Brasileiro de Genética, 2005, Águas de Lindóia.
51° Congresso Brasileiro de Genética. , 2005.

17. BECKER, D., JARDIM, F. M. A., WEBER, S. S., SANT'ANNA, F. H., SANTOS JR, S., RIBEIRO, M. O., SCHERER, L. C.
Contribuição do PCR in house no diagnóstico da tuberculose: a experiência em rotina de um laboratório de saúde pública In: XXIII Congresso Brasileiro de Microbiologia, 2005, Santos.
XXIII Congresso Brasileiro de Microbiologia. , 2005.

18. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Caracterização da região promotora do operon nifUSV de Azospirillum brasilense In: XXIV Reunião de Genética de Microorganismos, 2004, Gramado.
XXIV Reunião de Genética de Microorganismos. , 2004.

19. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Clonagem e caracterização dos genes glnB e glnD de Azospirillum amazonense In: XVI Salão de Iniciação Científica, Porto Alegre.
XVI Salão de Iniciação Científica. , 2004.

20. SANT'ANNA, F. H., SCHRANK, Irene Silveira
Determinação da atividade promotora do operon ORF2nifUSVORF4 de Azospirillum brasilense na presença da proteína NifA In: XV Salão de Iniciação Científica, 2003, Porto Alegre.
XV Salão de Iniciação Científica. , 2003.