

Universidade Federal do Rio de Janeiro, UFRJ

Estudos de genômica funcional em
Gluconacetobacter diazotrophicus:
ênfase no metabolismo redox e nas
vias alternativas de aminoacilação e
síntese de asparagina

Sylvia Maria Campbell Alquéres

Rio de Janeiro
2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Universidade Federal do Rio de Janeiro, UFRJ

Estudos de genômica funcional em
Gluconacetobacter diazotrophicus:
ênfase no metabolismo redox e nas
vias alternativas de aminoacilação e
síntese de asparagina

Sylvia Maria Campbell Alquéres

Orientador: Orlando B. Martins

Tese de Doutorado apresentada ao
Programa de Pós-Graduação em Química Biológica,
Instituto de Bioquímica Médica da UFRJ,
como parte dos requisitos necessários à
obtenção do título de Doutor em Química Biológica.

Rio de Janeiro
2010

Ficha catalográfica

Alquéres, Sylvia Maria Campbell A.

Estudos de genômica funcional em *Gluconacetobacter diazotrophicus*: ênfase no metabolismo redox e nas vias alternativas de aminoacilação e síntese de asparagina / Sylvia Maria Campbell Alquéres. Rio de Janeiro, 2010.

Tese (Doutorado em Química Biológica) –
Universidade Federal do Rio de Janeiro, Instituto
de Bioquímica Médica, 2010.

Orientador: Orlando Bonifácio Martins

1. *Gluconacetobacter diazotrophicus*. 2. Metabolismo redox 3. Síntese de asparagina. 4. Aminoacilação. 5. Fixação Biológica de Nitrogênio.

I. Martins, Orlando Bonifácio (Orient.). II.
Universidade Federal do Rio de Janeiro. Instituto
De Bioquímica Médica. III. Título.

Sylvia Maria Campbell Alquéres

Estudos de genômica funcional em *Gluconacetobacter diazotrophicus*: ênfase no metabolismo redox e nas vias alternativas de aminoacilação e síntese de asparagina

Orientador: Orlando Bonifácio Martins

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Química Biológica, Instituto de Bioquímica Médica da UFRJ, como parte dos requisitos necessários à obtenção do título de Doutor em Química Biológica.

Aprovada por:

Prof. Dr. Alexandre Rosado
Inst. Microbiologia CCS - UFRJ

Prof. Dr. Fábio Lopes Olivares
Centro de Biociências e Biotecnologia IB - UFRJ

Prof. Dr. José Ivo Baldani
Centro Nacional de Pesquisa de Agrobiologia - EMBRAPA

Prof. Dr. Russolina Zingali
Inst. Bioquímica Médica CCS - UFRJ

Prof. Dr. Orlando Martins
Inst. Bioquímica Médica CCS - UFRJ

Prof. Dr. Adriana Hemerly
Inst. Bioquímica Médica CCS - UFRJ

Dedicatória

À minha família: meus pais, irmãos e Tin. Ao Carlos Eduardo.

Agradecimentos

Ao Orlando. Já faz 10 anos que eu entrei no seu laboratório e em todo esse tempo eu devo ter pensado um milhão de vezes "uau, como eu sou sortuda de estar aqui". Além de gostar muito da maneira como você pensa as questões experimentais, da maneira com que você coloca os problemas científicos, eu ainda tive a oportunidade de ver o seu engajamento na luta por fazer UFRJ uma faculdade melhor, mais integrada, com mais recursos, que acolhesse mais alunos, e que em ultima instância, desempenhasse o seu papel de fazer do Rio uma cidade melhor.

Ao meu orientador e amigo Alex que sempre me ajudou muito em todos os trabalhos desenvolvidos, com bom humor e otimismo e uma leveza de se encarar os experimentos que sempre tornou o ambiente do laboratório mais divertido e entusiasmante.

Ao meu irmão Welington, que foi essencial para minha formação como bioquímica. Foi quem me explicou tudo sobre o então departamento de bioquímica médica, quem me deu aulas de metabolismo energético, me preparou para as provas do mestrado, e discutiu não só experimentos e mas também muitas coisas da vida.

Ao Zé, grande figura deste laboratório. Quero agradecer pela grande ajuda nos experimentos, mas principalmente na redação do paper, ao tomar a frente e fazer as coisas acontecerem quando eu estava demorando muito.

Ao professor Jonathan Trent, por abrir meus olhos para ver a poesia e o caos escondidos nos processos celulares.

Ao Eduardo Mattos, quem me explicou o básico de Glucona, quando eu não tinha nem idéia de por onde começar a estudar o assunto novo. Pelas críticas claras e essenciais ao trabalho.

Professor Ivo Baldani que me recebeu em seu laboratório onde passei ótimos mo-

mentos e tive resultados entusiasmantes. E é claro ao resto do grupo da Embrapa (Olavo inclusive), pelo excelente ambiente de trabalho.

A professora Adriana Hemerly pela revisão e acompanhamento desta tese e por abrir seu laboratório sempre que precisei fazer experimentos. E ao pessoal do JB, que me recebeu muito bem quando fui fazer experimentos lá.

Ao professor Paulo Ferreira.

Ao Marcelo Bertalan, por toda a ajuda com a bioinformática.

A Helma, que além de um doce é uma ótima companheira de trabalho, pela ajuda nos testes de ARA e discussões sobre os projetos.

Ao Luc, pela ajuda na construção dos mutantes aleatórios.

À Maria Clara, Milane e Marcos Sorgine, pela grande ajuda nos primeiros qPCRs.

Ao Jordano, pela ajuda na realização das HPLC.

Ao Marcus e a Renata, que durante o desenvolvimento da tese tiraram inúmeras dúvidas sobre sistema respiratório, mitocôndrias, ROS e etc. E deram ótimas dicas para o trabalho.

Ao Ricardo por me ensinar a importância de se fazer ciência usando uma pergunta e apenas duas amostras.

Ao Ramon e Nathalia, meus melhores amigos. Por estarem presentes nos principais momentos da minha vida. Por dividirem comigo uma visão original e clara sobre a vida e sobre as pessoas.

Ao professor Pedro pelo incentivo sempre.

Ao Rodrigo, meu orientador de mestrado que sempre foi um exemplo para mim.

Às meninas Aline, Cynthia, Vivi e Roberta, que trouxeram uma vida nova ao nosso laboratório.

À Tereza, Patrícia e Leonardo, da secretaria de pós graduação.

Resumo

A *Gluconacetobacter diazotrophicus* é um organismo endofítico estritamente aeróbico, isolado de cana-de-açúcar, que fixa nitrogênio (N_2). Seu sequenciamento foi finalizado em 2008 junto com os primeiros estudos de proteômica. Dentro deste cenário esse trabalho foi dedicado ao estudo da genômica funcional desta bactéria. Estes estudos foram focados em três diferentes temas: o primeiro deles foi a identificação dos transportadores presentes na bactéria pela análise de seu genoma, o segundo tema foi a influência de espécies reativas de oxigênio (ROS) no processo de fixação biológica de nitrogênio (FBN) por essa bactéria; e o terceiro foi a análise de uma via alternativa para a aminoacilação de tRNA e produção de asparagina.

Dentro do primeiro tema estudado observou-se que cerca de 45% dos transportadores anotados pertencem à classe de transportadores ativos primários, 30% à classe de transportadores ativos secundários e 12% são relacionados à formação de poros e canais na membrana celular. Entre os transportadores identificados vale ressaltar a presença de 23 genes transportadores do tipo TonB que estão ligados à captação de ferro extracelular, a presença de 40 genes da família RND e uma grande quantidade de genes para exportar metais pesados e outros compostos tóxicos, que podem conceder à bactéria uma notável resistência a ambientes inóspitos, e a presença de transportadores de um sistema de secreção do tipo IV, que confere à bactéria a capacidade de transportar macromoléculas através de seu envelope até a planta. Acredita-se que isso tenha um papel relevante no reconhecimento celular, no curso da interação planta-bactéria.

Dentro do segundo tema estudado abordamos a influência de ROS no processo de

FBN. Esse processo é catalizado pela nitrogenase e requer uma grande quantidade de ATP fornecida pela respiração aeróbica. No entanto, a atividade da nitrogenase é extremamente sensível à presença de oxigênio e possivelmente à ROS. Neste trabalho nós exploramos esse paradoxo e mostramos que de fato os níveis de ROS são diminuídos durante o processo de FBN devido ao aumento na expressão de seis enzimas detoxificadoras de ROS. Um agrupamento baseando-se nos padrões de expressão de cada um destes genes revelou a existência de um grupo formado pelos genes das enzimas superóxido dismutase, catalase do tipo E e subunidade MoFe da nitrogenase que apresenta 99,8% de similaridade de resposta nos padrões de expressão frente as diferentes condições apresentadas. Finalmente, a atividade da nitrogenase foi inibida de maneira dose-dependente por Paraquat, que aumenta os níveis de ROS intracelulares. Os resultados apresentados mostram que ROS inibe fortemente a atividade de nitrogenase e que a *G. diazotrophicus* muda o seu metabolismo redox durante a FBN aumentando a expressão e atividade de enzimas detoxificadoras de oxigênio que levam a uma menor quantidade de ROS gerada. Nós sugerimos que um forte controle da produção de ROS nesta fase crítica é um mecanismo adaptativo que permite a ocorrência do processo de FBN.

O terceiro tema aborda a presença e atividade de uma via alternativa para a produção de asparagina-tRNA (Asn-tRNA) e glutamina-tRNA (Gln-tRNA) em *G. diazotrophicus*. Essa via alternativa envolve a transaminação de um tRNA acilado erroneamente pela enzima codificada pelos genes *gatCAB*. A ausência de genes codificando para a asparagina sintetase e a falta de atividade desta enzima no extrato celular também indicam que o *gatCAB* pode ser responsável pela síntese deste aminoácido na célula. A influência da asparagina na expressão da nitrogenase também foi observada. Mutantes aleatórios com alterações no metabolismo de asparagina foram analisados e apresentaram taxas alteradas de FBN. Nós sugerimos que a existência de uma via alternativa para a biosíntese de asparagina pode ser um mecanismo adotado por alguns organismos para manter a regulação da FBN sob um controle maior.

Abstract

Gluconacetobacter diazotrophicus, is a nitrogen (N_2) fixating endophyte bacterium isolated from sugarcane, that is a strictly aerobic. The genome sequencing of this bacterium was completed in 2008, together with the first studies of proteomics. The work reported here was centered on the functional genomics of this bacteria. This study focuses on 3 different topics: the first concerns the identification of the transporters in the *G. diazotrophicus* genome; the second topic focuses on the influence of reactive oxygen species (ROS) in the biological nitrogen fixation (BNF) process; and the third topic focuses on the identification and analyses of an alternative route for aminoacylation and asparagine formation.

In the first topic we report that 45% of the transporters are primary active transporters, 30% are electrochemical potential-driven transporters (or secondary active transporters) and 12% are channel-type facilitators. Among these transporters are the 23 gene transporters of the TonB type involved in extracellular iron binding, the presence of 40 genes of the RND family and a large quantity of genes that can release heavy metals and other toxic composites thus providing the bacterium with resistance to inhospitable environments. Also noted were type IV secretion system transporters that enable the bacterium to transport macromolecules through its outer membrane to the plant. This probably has a relevant role in cell recognition during plant-bacterium interaction.

The second topic focuses on the influence of ROS in the BNF process. This process is catalyzed by a nitrogenase and requires copious amounts of ATP. Nitrogenase activity is extremely sensitive to inhibition by oxygen and reactive oxygen species (ROS). However, the high oxidative metabolic rates required to sustain biological nitrogen fixation (BNF)

may favor an increased production of ROS. Here, we explored this paradox and observed that ROS levels are, in fact, decreased in nitrogen-fixing cells due to the up regulation of transcript levels of six ROS detoxifying genes. A cluster analysis based on common expression patterns revealed the existence of a stable cluster with 99.8% similarity made up of the genes encoding the α -subunit of nitrogenase MoFe-protein, and the superoxide dismutase and catalase type E enzymes. Finally, nitrogenase activity was inhibited in a dose-dependent manner by paraquat, a redox cycler that increases cellular ROS levels. Our data revealed that ROS can strongly inhibit nitrogenase activity, and *G. diazotrophicus* alters its redox metabolism during BNF by increasing antioxidant transcript levels resulting in a lower ROS generation. We suggest that the carefully controlled ROS production during this critical phase is an adaptive mechanism to allow nitrogen fixation.

We also confirmed the presence of an alternative route for asparaginyl-tRNA (Asn-tRNA) and glutaminyl-tRNA (Gln-tRNA) formation. The alternative route involves transamidation of incorrectly charged tRNA via *gatCAB*. The absence from the *G. diazotrophicus* genome of genes encoding tRNA-independent asparagine synthetase and the lack of this enzyme in *G. diazotrophicus* extracts suggest that the *gatCAB* genes may be responsible for biosynthesis of asparagine in this organism. The influence of asparagine in the expression of nitrogenase gene was observed. Mutants with alterations in asparagine metabolism also presented different rates of biological nitrogen fixation compared to the wild type. We suggest that the presence of an alternative route to produce asparagine may give the cell a tighter control over BNF.

Sumário

Resumo	vi
Abstract	viii
Sumário	x
Lista de Figuras	xiv
Índice de Figuras	xx
Lista de Tabelas	xxi
Índice de Tabelas	xxii
Tabela de abreviações e fórmulas químicas	xxiii
1 Introdução	1
1.1 A Fixação de Nitrogênio: Origem e histórico.	3
1.2 A Fixação Biológica de Nitrogênio.	7
1.2.1 O complexo enzimático da nitrogenase	9
1.3 <i>Gluconacetobacter diazotrophicus</i>	11
1.3.1 Importância biotecnológica	11
1.3.2 Isolamento e habitat natural	14
1.3.3 Taxonomia	15
1.3.4 Características fisiológicas	16
2 Identificação e análise dos transportadores	19
2.1 Introdução	19
2.1.1 O sequenciamento de genomas	19
2.1.2 Anotação de genomas	23
2.1.3 Ferramentas de bioinformática	24
2.1.4 Transportadores e sistemas de transporte	25

2.2	Objetivos	30
2.3	Resultados e discussão	31
2.3.1	Visão geral dos transportadores.	42
2.3.2	Transportadores primários	44
2.3.3	Transportadores secundários	45
2.3.4	Poros e canais	47
2.4	Conclusões	48
3	As espécies reativas de oxigênio durante a FBN	50
3.1	Introdução	50
3.1.1	O paradoxo da FBN em organismos aeróbicos	50
3.1.2	A tolerância de <i>G. diazotrophicus</i> a O_2	51
3.1.3	Proteção respiratória	52
3.1.4	Ligação entre sistema respiratório e ROS	54
3.1.5	Enzimas detoxificadoras de ROS	55
3.1.6	A influência de ROS na FBN	56
3.1.7	Modulação de ROS durante a simbiose	57
3.1.8	Perspectivas	58
3.2	Objetivos	60
3.3	Resultados e Discussão	61
3.3.1	Visualização	61
3.3.2	Atividade Catalásica	63
3.3.3	Identificação dos genes responsáveis pela detoxificação de ROS no genoma de <i>G. diazotrophicus</i>	63
3.3.4	PCR em tempo real	67
3.3.5	Agrupamento	68
3.3.6	Fixação de nitrogênio na presença de paraquat	70
3.4	Conclusões	72
4	A via alternativa de aminoacilação de tRNA e o metabolismo de aspa- ragina	74
4.1	Introdução	74
4.1.1	A síntese proteica	74
4.1.2	A via alternativa de aminoacilação de tRNA.	77
4.1.3	A utilização da via alternativa para a produção de asparagina	81
4.1.4	Influência de asparagina e glutamina na FBN	82
4.2	Objetivos	84
4.3	Resultados e discussão	86

4.3.1	Localização e análise dos genes da via de aminoacilação de tRNA	86
4.3.2	Análise da expressão dos genes da via alternativa de síntese protéica	90
4.3.3	Influência de diferentes fontes de nitrogênio na expressão do gene da nitrogenase em função do tempo de crescimento	92
4.3.4	Utilização da via alternativa de acilação do tRNA para a formação de asparagina	97
4.3.5	Identificação da via alternativa em fixadores de nitrogênio	103
4.3.6	Screening dos mutantes auxotróficos para asparagina	104
4.3.7	Identificação e análise dos mutantes	105
4.4	Conclusões	114
5	Materiais e Métodos	116
5.1	Técnicas de bioinformática	116
5.1.1	Identificação de ORFs relacionadas ao sistema de transportadores no genoma da <i>G. diazotrophicus</i> PAL5	116
5.1.2	Análise dos clusters	119
5.2	Técnicas de biologia molecular	119
5.2.1	Extração de DNA genômico	119
5.2.2	Eletroforese de DNA	120
5.2.3	Extração do RNA total	120
5.2.4	Eletroforese de RNA	121
5.2.5	Síntese do cDNA	121
5.2.6	Amplificação do cDNA por reação em cadeia da polimerase (PCR)	122
5.2.7	PCR em tempo real ou PCR quantitativo (qPCR)	122
5.2.8	Cálculo da quantificação relativa	124
5.2.9	Obtenção dos mutantes aleatórios	126
5.2.10	Caracterização molecular dos mutantes aleatórios (PCR invertido)	127
5.2.11	Purificação de DNA para reação de sequenciamento	127
5.2.12	sequenciamento das regiões flanqueadores do transposon EZ-Tn5 TM <KAN- 2>Tnp transposome TM	127
5.3	Técnicas de microbiologia e bioquímica	128
5.3.1	Estirpes, meios e condições de cultivo	128
5.3.2	Determinação da produção de ROS em <i>G. diazotrophicus</i>	129
5.3.3	Atividade de catalase	130
5.3.4	Preparação de extrato celular livre de tRNA	130
5.3.5	Atividade de asparagina sintetase	130
5.3.6	Atividade de nitrogenase	131
5.3.7	Determinação de proteínas totais	131

5.3.8	Dosagem de aminoácidos livres por HPLC	132
5.3.9	Varredura dos mutantes auxotróficos para asparagina	134
6	Visão geral e perspectivas	135
	Referências Bibliográficas	137
	Apêndice	157

Lista de Figuras

1.1	Estrutura do complexo da nitrogenase de <i>Azotobacter vinelandii</i> . As subunidades Fe-proteína estão coloridas em azul claro, marron e cinza. As subunidades da MoFe-proteína estão coloridas em verde, amarelo, azul escuro e vermelho (Adaptado de Dixon [2004] por Guedes [2010]Oi).	11
2.1	Número genomas completos sequenciados até setembro de 2009 (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)	20
2.2	Número de projetos de sequenciamento de genoma depositados no GOLD™ (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)	21
2.3	Principal relevância de projetos de genoma bacterianos realizados no mundo (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)	22
3.1	Expressão relativa e atividade de redução do acetileno de células de <i>G. diazotrophicus</i> crescidas em meio LGI-P suplementado com 1 ou 20mM de sulfato de amônio.	61

-
- 3.2 Quantidade de ROS é menor durante FBN. Células de *G. diazotrophicus* crescidas em condições de fixação (FIX) e não-fixação de nitrogênio (NFIX) foram incubadas com o probe CM-H2DCFDA, sensível a ROS. As Figuras A, C, E e G representam o campo claro das imagens. As Figuras B, D, F e H são microscopias de fluorescência, e os pontos claros indicativos da presença de ROS. Todas as imagens foram obtidas com o mesmo tempo de exposição para permitir a comparação entre a intensidade dos sinais. Barra de escala: 10 μm 62
- 3.3 Atividade catalásica do extrato bruto de células de *G. diazotrophicus* crescidas em condições de fixação (FIX) e não-fixação de nitrogênio. A atividade catalásica dos extratos brutos celulares foi medida monitorando-se consumo de peróxido de hidrogênio a 240nm por 1 minuto à temperatura ambiente. Média +/- barra de erros de nove experimentos independentes são mostrados. 63
- 3.4 Expressão de enzimas de *G. diazotrophicus* envolvidas na detoxificação de ROS durante o processo de FBN. Expressão do mRNA correspondentes aos genes responsáveis pela detoxificação de ROS em células fixando e não-fixando nitrogênio crescidas durante 72h. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado pelo valor absoluto do gene na condição em que as células não fixam nitrogênio (NFIX). Os valores foram normalizados em relação a expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados. 67
- 3.5 Construção de clusters de expressão dos genes de detoxificação de ROS revela a presença de dois clusters principais: um formado pela *gorB* sozinha e outro formado por *gorA*, *kat*, *katC*, *nifD*, *sodA* e *katE*. Esse cluster pode ser dividido em três clusters menores: *kat* + *gorA*, *katC* e *katE* + *sodA* + *nifD*. 69
-

3.6	Efeitos de ROS no crescimento e na redução de acetileno ARA de células de <i>G. diazotrophicus</i> . A) Efeito de diferentes concentrações de paraquat (PQ) no crescimento de células fixando (em preto) e não-fixando (em cinza) nitrogênio. Depois de um período inicial de 24h, foi adicionado PQ para uma concentração final de 0,5 e 5 mM (linhas tracejadas e pontilhadas, respectivamente) (B) ARA de células depois de 72h de cultivo. A atividade foi medida depois de 48h na presença de 0,5 e 5mM de PQ.	70
4.1	Dogma central da Biologia molecular proposto por Francis Crick em 1958 .	74
4.2	Mecanismo geral de atuação das aminoacil-tRNA sintetases. Primeiramente a enzima se liga ao aminoácido e o liga à uma molécula de AMP, proveniente da clivagem de um ATP. Em seguida a enzima liga esse aminoacil ao seu tRNA correspondente e libera a molécula de AMP.	75
4.3	Mecanismo geral de síntese proteica.	76
4.4	Esquema e cartoon da via para a formação de $Asn-tRNA^{Asn}$ (similar para $Gln-tRNA^{Gln}$). Uma AspRS não-discriminativa forma $Asp-tRNA^{Asn}$. O aminoacil-tRNA formado é então corrigido por uma amidotransferase (<i>gatCAB</i>) na presença de ATP e de um doador do grupamento amida. . .	79

- 4.5 Alinhamento da porção N-terminal de algumas aspartyl-tRNA sintetases bacterianas não discriminativas. A porção azul, indica a posição correspondente à Prolina 77 de arqueas. Em bactérias, essa é uma posição conservada, onde encontra-se sempre uma arginina, independente de ser uma aa-tRNS sintetase discriminativa ou não. A posição vermelha, correspondente à asparagina (N) de *E. coli*, em enzimas não discriminativas encontramos sempre Treonina (T), Leucina (L) ou Alanina (A). Já na porção marcada de amarelo (M87 de *E. coli*) a presença de uma leucina (L) diminui fortemente a especificidade por *tRNA^{Asp}* e aumenta para *tRNA^{Asn}*. Foram usadas as seqüências dos seguintes organismos: *Bordetella pertussis* Tohama I, *Escherichia coli* str. K-12 substr. W3110, *Helicobacter pylori* 26695, *Gluconacetobacter diazotrophicus* PAI 5, *Brucella melitensis* bv. 1 str. 16M. 89
- 4.6 Localização do *gatCAB* no genoma da *G. diazotrophicus*. 90
- 4.7 Expressão de enzimas da via alternativa de aminoacilação de tRNA em função da disponibilidade de nitrogênio - células crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIK) de sulfato de amônio - após 72h de crescimento. A expressão dos genes na condição de FIX de crescimento foi utilizada como calibradora e, portanto, fixada em 1. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado (Ct) pelo valor absoluto do gene na condição calibradora. Os valores foram normalizados em relação à expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados. 91
- 4.8 Crescimento de *G. diazotrophicus* em meio LGI-P líquido suplementado com 1 ou 20mM de sulfato de amônio ou suplementado com 20mM de asparagina, aspartato, glutamina ou glutamato. As células foram cultivadas a 30°C, 150 rpm. 93

- 4.9 Expressão do gene *nifD* em função da disponibilidade de N no meio de cultura. As células foram crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIX) de sulfato de amônio - e a expressão relativa do gene *nifD* foi verificada. A expressão do gene na condição FIX com 24h de cultivo foi utilizada como calibradora e, portanto, fixada em 1. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado (Ct) pelo valor absoluto do gene na condição calibradora. Os valores foram normalizados em relação à expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados. 96
- 4.10 Relação entre os níveis intracelulares de asparagina, glutamina e aspartato (apresentados na Tabela 4.3) e a expressão do *nifD* (apresentados na Figura 4.9). Podemos observar no primeiro gráfico que enquanto os níveis de asparagina estão até em torno de 40 uM, ainda pode-se observar a expressão do gene *nifD*. A partir desta concentração a expressão do gene *nifD* encontra-se sempre reprimida. A análise com glutamina e aspartato não mostra nenhuma relação entre as concentrações intracelulares destes aminoácidos e a expressão do gene *nifD*. 97
- 4.11 Atividade de asparagina sintetase dependente de tRNA. Quantidade de [¹⁴C]asparagina formada pela conversão de [¹⁴C]aspartato por células de *G. diazotrophicis* e *E. coli* (controle positivo) livres de tRNA. Média +/- barra de erros de três experimentos independentes são mostrados. 102
- 4.12 Exemplo de placas utilizadas para a seleção dos mutantes auxotróficos para asparagina. Os mutantes que não cresceram na placa “A” (meio LGI-P suplementado com sulfato de amônio), mas cresceram na placa “B” (meio LGI-P suplementado com asparagina) foram selecionadas para experimentos subseqüentes. 105

-
- 4.13 Representação gráfica da via metabólica do ciclo do ácido tricarboxílico, onde está destacado (em vermelho) a enzima PEPCase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes. 107
- 4.14 Atividade de nitrogenase (em nmoles de etileno . mg^{-1} . h^{-1}) da estirpe selvagem PAL5 e mutante fosfoenolpiruvato carboxilase (MUT4) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados. 108
- 4.15 Representação gráfica via do metabolismo de glutamato, onde destaca-se (em vermelho) a enzima Amidofosforibosil transferase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes. 109
- 4.16 Atividade de nitrogenase (em nmoles de etileno . mg^{-1} . h^{-1}) da estirpe selvagem PAL5 e mutante Amidofosforibosil transferase (MUT14) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados. 110
- 4.17 Atividade de nitrogenase (em nmoles de etileno . mg^{-1} . h^{-1}) da estirpe selvagem PAL5 e mutante Permease do sistema de transporte de D-metionina metI (MUT98) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados. 111
- 4.18 Representação gráfica da via de metabolismo de nitrogênio, onde destaca-se (em vermelho) a enzima Glutamato sintase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes. 112

- 4.19 Atividade de nitrogenase (em nmoles de etileno . mg^{-1} . h^{-1}) da estirpe selvagem PAL5 e mutante Glutamato sintase (MUT35) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados. 113
- 6.1 Esquema mostrando como a obtenção de determinados genes conferiu vantagens à *G. diazotrophicus* e a permitiu colonizar um novo nicho ecológico. 136

Lista de Tabelas

1.1	Exemplos de gêneros de organismos diazotróficos	8
2.1	Descrição das ORFs identificadas com pertencentes ou relacionadas com o sistema de transportes no genoma da <i>G. diazotrophicus</i>	32
2.2	Comparação entre o sistema de transportes no genoma da <i>G. diazotrophicus</i> com outros genomas publicados	43
2.3	Principais classes de transportadores anotados no genoma de <i>G. diazotrophicus</i>	44
3.1	Identificação dos genes responsáveis por detoxificação de ROS em <i>G. diazotrophicus</i>	64
3.2	Tipos de catalase	66
4.1	Influência de diferentes fontes de nitrogênio na atividade da nitrogenase, utilizando-se o método de redução do acetileno, em <i>G. diazotrophicus</i> crescida em meio LGI-P a 30°C após 72h de crescimento.	83
4.2	Influência de diferentes fontes de nitrogênio na expressão do gene da nitrogenase em função do tempo de crescimento	94
4.3	Concentração em μM medida por HPLC de aminoácidos livres intracelulares de células de <i>G. diazotrophicus</i> crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIX) de NH_4SO_4 em função do tempo de cultivo.	95

4.4	Identificação da enzima gatCAB, da via alternativa de acilação de tRNA, em todos os organismos que não possuem os genes que codificam para as enzimas asparagina sintetase A ou B (<i>asnA</i> e <i>asnB</i>).	99
4.5	Identificação da via de síntese de asparagina e da via alternativa de acilação de tRNA em organismos com capacidade de fixar nitrogênio.	104
4.6	Seqüências obtidas dos produtos de iPCR dos mutantes aleatórios.	106
5.1	Primers qPCR	124
5.2	Eficiência dos oligonucleotídeos. Nome do gene, temperatura de desnaturação (TM), o R^2 da curva de eficiência e a eficiência de amplificação de cada um dos oligonucleotídeos utilizados nesse trabalho. A eficiência foi calculada utilizando a fórmula $E = 10^{(-1/slope)}$ e $\%E = (E - 1) * 100$	126
5.3	Quantidades de ágar acrescidos aos meios líquidos a fim de se obter meios semi-sólidos ou sólidos.	129

Tabela de abreviações e fórmulas químicas

23S	DNA que codifica para o gene 23S
aaRS	aminoacil-tRNA sintetase
ABC	“ATP-binding Cassette Transporter”
ADP	Adenosina Difosfato
AdT	Amidotransferase tRNA específica
AMP	Adenosina monofosfato
ARA	Atividade de redução de acetileno
Asn	Asparagina
Asp	Aspartato
ATP	Adenosina Trifosfato
BASF	Badische Anilin- und Soda-Fabrik
BLAST	“Basic Local Alignment Sequence Tool”
cDNA	DNA complementar
COG	”Clusters of Orthologous Groups”
Ct	“Threshold cycle”
D-AspRS	Aspartil-tRNA sintetase específica
DNA	Ácido Desoxirribonucleico
DTT	Ditiotreitol
EC	“Enzyme Comission”

EDTA	Ácido etilenodiamino tetra-acético
EMBL	“European Molecular Biology Laboratories”
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
FAPERJ	Fundação de Amparo a Pesquisa do Estado do Rio de Janeiro
FBN	Fixação Biológica de Nitrogênio
Fe-proteína	Ferro-proteína da nitrogenase
FIX	células crescidas com 1 mM de sulfato de amônio
g	Gramas
g/L	Gramas por Litro
GDB	“Human Genome Database”
GDI	<i>Gluconacetobacter diazotrophicus</i>
GLIMMER	“Gene Locator and Interpolated Markov ModelER”
Gln	Glutamina
Glu	Glutamato
Glu/Asp-AdT	Glutamina/asparagina amidotransferase tRNA específica
GOGAT	Glutamato sintase
GOLD TM	Genomes OnLine Database
GR	Glutathione redutase
GSDB	“Genome Sequence Data Base”
GTP	Guanosine-5'-triphosphate
h	Hora
HPLC	Cromatografia líquida de alto desempenho
IAA	Ácido indol acético
iPCR	PCR invertido
Kat	Catalase
<i>KCl</i>	Cloreto de potássio

KEGG	”Kyoto Encyclopedia of Genes and Genomes”
kPa	Kilopascal
LNCC	Laboratório Nacional de Computação Científica
Mb	Milhões de pares de bases
MFS	“Major Facilitator Superfamily”
mL	Mililitro
mM	Milimolar
MoFe- proteína	Proteína Ferro-Molibdênio da nitrogenase
<i>NaOH</i>	Hidróxido de sódio
NCBI	”National Center for Biotechnology Information”
ND-AspRS	Aspartil-tRNA sintetase não discriminativa
NFIX	células crescidas com 20 mM de sulfato de amônio
<i>NH₃</i>	Amônia
<i>NH₄</i>	Amônio
<i>(NH₄)₂SO₄</i>	Sulfato de amônio
nmoles	nanomoles
C	Celsius
ORF	Fase de leitura aberta
PCR	Reação em cadeia da polimerase
PDB	“Protein Data Bank”
PEPCase	Fosfoenol piruvato carboxilase
pH	Potencial Hidrogeniônico
pHi	Potencial Hidrogeniônico interno
pHo	Potencial Hidrogeniônico externo
PHYLIP	“Phylogeny Inference Package”
PIR	“Protein Information Resource”
PMSF	Fenilmetilsulfonil fluorido

PQ	Paraquat
qPCR	PCR quantitativo
RNA	Ácido ribonucleico
RND	“Resistance Nodulation Division”
ROS	Espécies reativas de oxigênio
rRNA	RNA ribossômico
SOD	Superóxido dismutase
TAE	Tris Acetato EDTA
TCA	Ácido tricloroacético
TCDB	“Transport Classification Database”
TIGR	The Institute for Genomic Research
TM	Marca registrada
tRNA	RNA transportador ou ARN de transferência
UENF	Universidade Estadual do Norte Fluminense
UERJ	Universidade Estadual do Rio de Janeiro
UFRJ	Universidade Federal do Rio de Janeiro
w/v	massa por volume
WGS	“Whole-Genome Shotgun Sequencing”
µg	Micrograma
µM	Micromolar

Capítulo 1

Introdução

O desenvolvimento do equipamento e das técnicas que permitem o sequenciamento automatizado de fragmentos de DNA e de genomas completos está na base do aparecimento da nova área científica da Genômica. De fato, desde 1995, data em que foi disponibilizada a primeira sequência completa do genoma de um ser vivo (da espécie bacteriana *Haemophilus influenza*), as abordagens genômicas e pós-genômicas vieram revolucionar a forma de investigar e de desenvolver qualquer atividade nas várias vertentes da Biologia. De acordo com o banco de dados GOLDTM - Genomes OnLine Database v. 3.0 (www.genomesonline.org) - existem hoje, março de 2010, cerca de 1100 genomas completos publicados, incluindo o genoma humano, e acredita-se que até o final de 2010, cerca de 5000 sequências estarão disponíveis.

Esta quantidade de informação genômica abriu as portas para o surgimento da proteômica, que é o estudo do conjunto das proteínas expressas em uma célula. Veio também incentivar a nova área inter- e trans-disciplinar da Bioinformática, que envolve a Biologia, a Estatística e a Informática. As informações geradas por essas abordagens permitiram a construção de numerosas bases de dados, com informação em escala genômica e ferramentas computacionais associadas.

O grande desafio dos pesquisadores que atualmente trabalham nestas áreas de conhecimento é conseguir relacionar a grande quantidade de informação gerada pelas análises

genômicas e proteômicas com os aspectos fisiológicos do organismo estudado, entender finalmente como os genes afetam a fisiologia das células e são por ela afetados. A seu favor, aqueles que se aventuram por essa área contam com essa enorme quantidade de informação gerada por análises “high-throughput”. Por outro lado, essa grande quantidade de informação gera uma grande quantidade de perguntas e cabe ao pesquisador selecionar aquelas mais relevantes para a compreensão da fisiologia do organismo estudado.

Dentro deste cenário esse trabalho foi dedicado ao estudo da genômica funcional da bactéria *Gluconacetobacter diazotrophicus*. Seu sequenciamento foi finalizado em 2008 junto com os primeiros estudos de proteômica. Devido à complexidade de seu genoma e de suas interações com a cana-de-açúcar, seu principal hospedeiro, essa análise funcional requer o conhecimento da fisiologia integrativa. No presente trabalho utilizamos diversas técnicas que vão desde a análise da expressão gênica utilizando a técnica de PCR em tempo real (qPCR, de quantitative polymerase chain reaction), até a quantificação de intermediários metabólicos por cromatografia, passando por técnicas de microscopia de fluorescência e enzimologia. Estes estudos foram focados principalmente em aspectos da fixação biológica de nitrogênio e síntese de proteínas e na habilidade da bactéria em fixar nitrogênio na presença de oxigênio. Apesar de voltados para uma possível aplicação biotecnológica, são estudos que abrem novas perspectivas de compreensão de mecanismos fisiológicos do organismo estudado e que podem vir a ser utilizados por outros modelos de estudo.

Devido a variedade de temas abordados esta tese foi separada em seis capítulos:

- O **primeiro capítulo (1)** consiste numa introdução geral sobre a fixação de nitrogênio e sobre o modelo de estudo utilizado.
- O **segundo capítulo (2)** aborda o processo de identificação e análise dos transportadores de *G. diazotrophicus*.
- No **terceiro capítulo (3)** se encontra um estudo sobre a influência de espécies

reativas de oxigênio (ROS) no processo de FBN e do estado redox da célula durante este processo.

- O **quarto capítulo** (4) trata da existência de uma provável via alternativa de aminoacilação de tRNA em *G. diazotrophicus*, sua influência na formação de aspargina e suas implicações no processo de FBN.
- O **quinto capítulo** (5) apresenta os materiais e métodos utilizados para a realização dos estudos tratados nos capítulos anteriores.
- O **sexto capítulo** (6) apresenta uma conclusão geral da tese e perspectivas deste trabalho e é seguido pelas **Referências bibliográficas** e pelo **Apêndice**, que contém os manuscritos dos trabalhos desenvolvidos durante esta tese.

1.1 A Fixação de Nitrogênio: Origem e histórico.

Formas complexas e variadas de vida existem na Terra há pelo menos 3,5 bilhões de anos, e, independente de suas diferenças entre si, todas são fundamentadas na existência de macromoléculas biológicas. Estas macromoléculas podem ser divididas em quatro grandes grupos: ácidos nucleicos, proteínas, carboidratos e lipídios. O nitrogênio é o elemento básico de dois destes grupos de macromoléculas: os ácidos nucleicos e as proteínas. Embora amplamente presente na sua forma orgânica, a maior parte do nitrogênio da Terra está na sua forma gasosa (N_2). A produção da forma básica de nitrogênio orgânico, a amônia (NH_3) requer necessariamente a quebra do nitrogênio atmosférico e sua ligação com átomos de hidrogênio (H). O processo de transformação do nitrogênio atmosférico na sua forma orgânica (NH_3) é chamado então de fixação de nitrogênio. Esta reação, embora seja termodinamicamente favorável, não ocorre espontaneamente visto que os dois átomos de nitrogênio da molécula de N_2 encontram-se unidos através de uma tripla ligação muito estável e necessitam, para o rompimento destas, uma energia de ativação muito

alta. Considerando-se a importância da forma orgânica do nitrogênio para a formação de macromoléculas essenciais para a vida, o estudo do processo de fixação de nitrogênio recebeu grande atenção desde o final do século XVIII. Foi nesta época que descobriu-se que a amônia era composta de um átomo de nitrogênio e três de hidrogênio. A partir de tal descoberta, químicos do mundo todo tentavam sintetizar amônia a partir destes gases, abundantes na natureza, sem sucesso. Foi o químico Fritz Haber que em 1909 chegou à solução deste problema de longa data na química. Quando ele e seu assistente Robert Le Rossignol o conseguiram em difíceis condições de laboratório, o processo produzia apenas uma pequena quantidade de amônia em longos espaços de tempo, excluindo qualquer esperança de produção industrial. Era preciso descobrir um catalisador para acelerar a reação. Após experiências com incontáveis metais, descobriu que o pó do raro metal ósmio (do qual existiam apenas 110 quilos no mundo) produzia o efeito desejado. Num dos grandes dias da química, 2 de julho de 1909, Haber fez uma demonstração de sua produção de amônia, a uma taxa de setenta gotas por minuto, na presença de dois diretores técnicos de uma grande empresa química (Badische Anilin- und Soda-Fabrik, BASF), Alywin Mittasch e Carl Bosch. A BASF pressionou então Bosch e Mittasch a melhorar o processo de Haber. Embora a empresa tivesse comprado todo o estoque de ósmio do mundo, a equipe estava decidida a encontrar um catalisador mais rápido e mais abundante. Após quatro mil experiências, produziram um catalisador ideal, composto de ferro, óxidos de alumínio, cálcio e potássio. A receita continua praticamente sem mudança até hoje. O processo Haber-Bosch, como veio a ser conhecido, responde por 99% do nitrogênio orgânico produzido no mundo, o equivalente a cerca de 130 milhões de toneladas de amônia por ano; 4/5 disso vão para os fertilizantes. Como o nitrogênio fixado quase sempre foi o maior limitante na agricultura, a descoberta do processo de fixação de nitrogênio representou um enorme avanço para a humanidade e permitiu que esta alcançasse a incrível marca 6,6 bilhões bilhões de habitantes em 2007, em vez dos 3,6 bilhões que seriam o máximo previsto sem essa descoberta [Goran, 1947].

Apesar da grande eficiência deste processo, ele ainda necessita empregar altas tem-

peraturas (400 a 600°C), pressões elevadas (100 a 200 atm) e utiliza catalisadores a base de ferro. Isso representa um gasto enorme de fontes energéticas não renováveis que geralmente é calculado em seis barris de petróleo por tonelada de NH_3 sintetizada. Considerando que a limitação de nitrogênio é um dos principais problemas na produtividade agrícola e que a utilização de fertilizantes nitrogenados é largamente utilizada, pode-se calcular que o gasto energético utilizado na produção destes fertilizantes chega à incrível marca de 5% da produção mundial de gás natural [Smil, 2000]. Além disso, a larga utilização de fertilizantes nitrogenados acabou com a necessidade da utilização do tradicional meio de enriquecimento do solo por rotatividade das culturas e da utilização de esterco animal para fertilizar o solo. O resultado disso é uma contínua perda de matéria orgânica do solo, levando-se em última instância a um desequilíbrio do solo e a perda da sua capacidade de reter água e manter populações de microrganismos. Atualmente, a maioria dos campos agrícolas não funciona como unidades recicladoras da matéria orgânica gerada. Observa-se, ao contrário, todo o rejeito animal e de plantas, que são ricos em nitrogênio, sendo dispensados como lixo comum causando um enorme impacto ambiental. Como agravante, tal problema ocorre principalmente em locais onde a prática inapropriada do cultivo agrícola, incentivado pelos baixos preços dos fertilizantes nitrogenados que, aliado à política de subsídios agrícolas, aumentam ainda mais o impacto de tais medidas. Nessas circunstâncias apenas metade (no máximo) do nitrogênio aplicado no solo chega às plantas. O resto é lixiviado pela chuva e atinge lençóis freáticos, rios e águas costeiras. O resultado é a crescente eutrofização de tais ambientes, altamente prejudicial à vida marinha. E há de se considerar ainda o perigo para a saúde humana quando tais lençóis freáticos entram em contato com fontes de água potável.

Além de todos estes problemas, a oxidação da uréia e do amônio à nitrato progressivamente levam à acidificação do solo – um dos principais problemas agrícolas no mundo. Outro problema relacionado é a emissão de óxidos do nitrogênio (NO_x) por microrganismos agindo nos fertilizantes. Estes óxidos são altamente voláteis, o que causa a perda de quase 50% da amônia aplicada, além disto, estes óxidos depois de volatilizados acabam

voltando para o solo, desta vez de uma maneira mais drástica, através da chuva ácida.

Alguns destes problemas podem ser contornados simplesmente pelo uso mais eficiente dos fertilizantes aplicados, utilizando-se doses adequadas e respeitando-se os tempos entre as aplicações. Uma outra abordagem que contorna os problemas descritos é a utilização da fixação biológica de nitrogênio (FBN). A FBN é um processo realizado por diferentes classes de microrganismos, exclusivamente procarióticos, que consiste na transformação do nitrogênio atmosférico em amônia através de uma enzima chamada de nitrogenase. Do ponto de vista energético, ele também é dispendioso para o microrganismo que o realiza, porém, por ser catalizado por uma enzima, a energia de ativação necessária para a ocorrência do processo é bem menor do que no processo inorgânico. De fato, o próprio Fritz Haber chamou a atenção para a FBN quando terminou seu discurso de aceitação do Nobel de Química, em 1920, com as seguintes palavras:

“Pode ser que esta não seja a solução final. As bactérias do nitrogênio nos dizem que a Natureza, com suas sofisticadas formas de química da matéria viva, compreende e utiliza métodos que nós ainda não sabemos imitar. Neste meio tempo, nos basta que a indústria química traga as riquezas nutritivas e venha ajudar o camponês que, na boa terra, transforma pedras em pão.”¹

Quatro anos depois deste discurso de Haber, em Aussig, na República Tcheca, nasceu Johanna Döbereiner. Em 1950 ela se formou em agronomia pela Universidade de Munique apresentando monografia de conclusão de curso intitulada “Bactérias na fixação assimbiótica de nitrogênio e a possibilidade de seu aproveitamento na agricultura”. Poucos meses depois emigrou para o Brasil, onde foi contratada pelo então Instituto de Ecologia e Experimentação Agrícola, atual Centro Nacional de Pesquisa de Agrobiologia da Embrapa, localizado no município de Seropédica, estado do Rio de Janeiro. Brasileira naturalizada, publicou uma série de trabalhos sobre o enriquecimento seletivo de bactérias fixadoras de nitrogênio. Os grupos que dirigiu na Embrapa e na Universidade Federal Rural do Rio de Janeiro iniciaram, em 1963, um extenso programa de pesquisas sobre vários aspectos da fixação biológica do nitrogênio por plantas cultivadas, acumulando dados e resultados

¹Tradução livre de Fritz Haber, Haber, F. 1920. The synthesis of ammonia from its elements. http://nobelprize.org/nobel_prizes/chemistry/laureates/1918/haber-lecture.pdf

que indicam a superioridade desses recursos naturais sobre a utilização de fertilizantes minerais.

Na introdução oficial do cultivo da soja no Brasil, no final da década de 50, Döbereiner se posicionou a favor do aproveitamento das associações entre a planta e bactérias fixadoras de nitrogênio, opondo-se à utilização obrigatória de adubos nitrogenados. A adoção desta linha de pensamento resultou, ao longo dos anos seguintes, numa considerável economia para o país. Seus estudos sobre fixação biológica de nitrogênio proporcionaram o desenvolvimento de uma tecnologia capaz de diminuir ou até mesmo eliminar a dependência de adubação nitrogenada na cultura, poupando atualmente ao país entre um e dois bilhões de dólares por ano. Tal tecnologia faz com que o Brasil tenha o menor custo de produção de soja do mundo, se estabelecendo como o segundo maior produtor mundial do grão. As possibilidades abertas pelo achado em relação à atividade agrícola motivaram a criação do Programa de Cooperação Internacional em Fixação de Nitrogênio nos Trópicos, sob coordenação de Johanna Döbereiner. Em 1997, assim como Fritz Haber em 1918, ela foi indicada para o Prêmio Nobel de Química por seus trabalhos sobre a fixação biológica de nitrogênio.

1.2 A Fixação Biológica de Nitrogênio.

O processo de FBN é realizado por diferentes classes de microrganismos, chamados de microrganismos diazotróficos. Eles são exclusivamente procarióticos, mas podem ser arqueas metanogênicas ou bactérias aeróbicas, anaeróbicas, Gram positivas, Gram negativas, heterotróficas, autotróficas e fototróficas.

Tabela 1.1: Exemplos de gêneros de organismos diazotróficos

De vida livre		Simbióticos		Endofíticos
Aeróbicos	Anaeróbicos	Leguminosas	Não-leguminosas	
<i>Azotobacter</i>	<i>Firmibacterias</i>	<i>Rhizobium</i>	<i>Frankia</i>	<i>Azospirillum</i>
<i>Beijerinckia</i>	<i>Desulfovibrio</i>	<i>Bradyrhizobium</i>	<i>Anabaena</i>	<i>Gluconacetobacter</i>
<i>Nostoc</i>	<i>Rhodospirillum</i>			<i>Azoarcus</i>
	<i>Methanococcus</i>			<i>Herbaspirillum</i>

Os organismos diazotróficos são divididos em três grupos principais:

- os de **vida livre**, que podem viver e fixar N_2 independentemente de outros organismos. A capacidade de fixar N_2 por microorganismos de vida livre está amplamente distribuída entre as bactérias e arqueas do solo e da água. Todos os grupos fisiológicos tem algum representante que é capaz de realizar este processo com maior ou menor eficiência [Baldani et al., 1997].
- os que **vivem em simbiose** com outros organismos. Os exemplos mais conhecidos fazem parte dos gêneros *Rhizobium* e *Bradyrhizobium*, que infectam as raízes de plantas leguminosas e estimulam a produção de nódulos, dentro dos quais as bactérias fixam N_2 , enquanto a planta responde fornecendo nutrientes orgânicos produzidos durante a fotossíntese. Tais nódulos possuem condições otimizadas para uma máxima eficiência do processo de fixação. Outra bactéria fixadora de N_2 que estabelece relações simbióticas (porém com plantas não-leguminosas) é a cianobactéria *Anabaena azollae*, que vive em pequenos poros da planta aquática *Azolla*. Outro exemplo é o o gênero *Frankia*, uma bactéria que se assemelha a *Streptomicetes*, que forma nódulos em árvores, arbustos e herbáceas [Baldani et al., 1997].
- A interação entre bactérias diazotróficas e gramíneas é um tipo particular de associação entre planta/diazotróficos conhecido como colonização **endofítica**. A associação endofítica é caracterizada por uma colonização de diversos tecidos vegetais por bactérias diazotróficas que promovem o crescimento da planta sem causar qualquer sintoma de doença [Döbereiner, 1992]. Baldani e colaboradores sugeriram a

separação do grupo entre os "endofíticos associativos ou facultativos" para os microorganismos que são capazes de colonizar a superfície e o interior da raiz e sobreviver bem no solo, e os "endofíticos obrigatórios" que não sobrevivem no solo mas colonizam o interior das raízes e a parte aérea da planta [Baldani et al., 1997]. Há, porém, outros autores que discordam quanto ao uso da terminologia "endofítico obrigatório" pois estas bactérias podem crescer em meio de cultura sem a presença de extratos de vegetais [Reinhold-Hurek e Hurek, 1998]. Mais recentemente, devido a dificuldade de separação entre esses dois grupos esta classificação foi abandonada. A lista de diazotróficos endofíticos que colonizam plantas não-leguminosas inclui várias bactérias do gênero *Azospirillum* (*A. brasilense*, *A. lipoferum*, *A. amazonense*, *A. irakense*), *Herbaspirillum seropedicae*, *H. rubrisubalbicans*, *G. diazotrophicus*, *Azoarcus sp.* e *Burkholderia sp.* Estas bactérias foram isoladas de cana-de-açúcar, plantas forrageiras, cereais e outros hospedeiros e são descritas em detalhes em publicações de Döbereiner e Baldani [Döbereiner et al., 1993, Baldani et al., 1997].

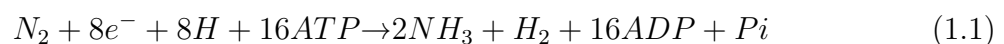
Organismos fixadores de nitrogênio que estabeleceram simbiose com plantas superiores sempre foram o foco da atenção dos pesquisadores que estudam o processo de FBN. De fato, as maiores contribuições em nitrogênio fixado por FBN foram detectadas em oceanos e plantas leguminosas. No entanto, evidências recentes de significativa FBN em gramíneas de importância econômica, em especial cana-de-açúcar (*Saccharum sp.*), arroz (*Oryza sativa*), gramíneas forrageiras, como *Leptochloa fusca* despertaram novos interesses nos estudos de fixação de nitrogênio e o estudo da FBN em gramíneas ganhou novo fôlego [Baldani et al., 1997, James, 2000].

1.2.1 O complexo enzimático da nitrogenase

Em organismos diazotróficos, o processo de redução do nitrogênio é catalizado pela nitrogenase, que é um complexo de metaloenzimas com características mecânicas e estruturais conservadas [Rees e Howard, 2000, Lawson e Smith, 2002, Dixon e Kahn, 2004]. Estas enzimas contêm dois componentes que são nomeados de acordo com o metal de

sua composição. O componente dimérico menor, conhecido como Ferro-proteína (Fe-proteína, produto do gene *nifH*), que funciona como um doador de elétrons dependente de ATP para o maior componente heterotetramérico, conhecido como proteína Ferro-molibdênio (MoFe-proteína, produto do gene *nifDK*), que contém o sítio catalítico da enzima [Dixon e Kahn, 2004]. O complexo da nitrogenase é codificado pelos genes estruturais *nifHDK*. A Ferro-proteína contém um núcleo 4Fe-4S, já a MoFe-proteína (codificada pelos genes *nifDK*), contém dois grupos prostéticos: Centro P [8Fe-7S] e o cofator ferro molibdênio. Estes componentes apresentam forte homologia entre diferentes microorganismos diazotróficos, podendo o componente I de um organismo ser combinado ao componente II de outro, tornando o complexo nitrogenase ativo [Detroy et al., 1967].

Um modelo estrutural do complexo nitrogenase de *Azotobacter vinelandii* está apresentado na Figura 1.1 Todos os diazotróficos têm um sistema ferro-molibdênio nitrogenase, mas em condições de ausência de molibdênio, alguns organismos como *A. vinelandii* e *Rhodobacter capsulatus* - induzem a síntese de nitrogenases alternativas contendo como co-fatores o ferro-vanádio ou ferro-ferro [Eady, 1996]. O mecanismo enzimático da nitrogenase requer a redução da Fe-proteína por doadores de elétrons como ferredoxina e flavodoxina, a transferência de elétrons para a MoFe-proteína através de um processo dependente da hidrólise de Mg-ATP e a transferência de elétrons para o substrato ligado ao sítio ativo da MoFe-proteína [Dixon e Kahn, 2004]. Em condições ótimas, a estequiometria global da reação é [Simpson e Burris, 1984]:



Devido à sensibilidade das nitrogenases ao oxigênio e ao alto custo energético da FBN, esse processo em diazotróficos é estritamente regulado [Postgate, 1998]. Em proteobactérias, a síntese do complexo enzimático da nitrogenase depende da atividade da proteína *NifA* em conjunto com a o promotor σ^{54} responsável pela ativação da transcrição dos genes *nif* (genes da fixação do nitrogênio), que codificam a nitrogenase e outras proteínas necessárias para a síntese desta enzima, maturação dos seus componentes, transporte

de elétrons e regulação da transcrição [Merrick, 1992].

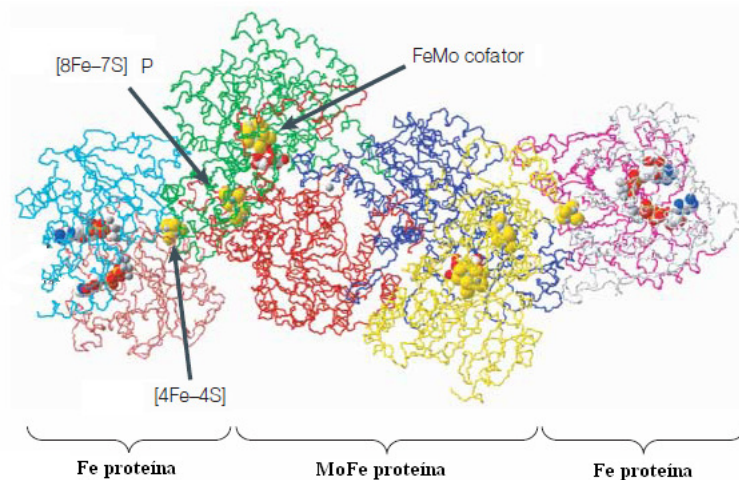


Figura 1.1: Estrutura do complexo da nitrogenase de *Azotobacter vinelandii*. As subunidades Fe-proteína estão coloridas em azul claro, marron e cinza. As subunidades da MoFe-proteína estão coloridas em verde, amarelo, azul escuro e vermelho (Adaptado de Dixon [2004] por Guedes [2010]Oi).

1.3 *Gluconacetobacter diazotrophicus*

1.3.1 Importância biotecnológica

As maiores contribuições da FBN são detectadas em oceanos (devido a ação das cianobactérias) e em plantas leguminosas (devido a ação das bactérias simbióticas), mas atualmente grandes esforços têm sido feitos para que essa fixação seja estendida para outras espécies que colonizam plantas não leguminosas, principalmente gramíneas de grande importância econômica. No caso do Brasil a cultura de cana-de-açúcar (*Saccharum spp.*; família *Poaceae*) é considerada estratégica no desenvolvimento, graças ao PróÁlcool, um programa que visa incentivar o uso do etanol alternativamente a combustíveis fósseis [Zanin et al., 2000]. Além disso, o Brasil é o maior produtor mundial de cana-de-açúcar segundo o Instituto Brasileiro de Geografia e Estatística (IBGE). Neste cenário é necessário o desenvolvimento de tecnologias capazes de diminuir o custo de produção da cana-de-açúcar e que permitam um aumento da produtividade das áreas plantadas. A descoberta

de que a associação de plantas da família *Poacea* (*Gramineae*) com algumas bactérias diazotróficas era capaz de promover o crescimento vegetal [Baldani et al., 1997] estimulou as pesquisas nesta área. Entre essas pesquisas, resultados bastante promissores foram obtidos em plantas de cana-de-açúcar inoculadas com *G. diazotrophicus*. A inoculação de plantas de cana-de-açúcar micropropagadas com a estirpe Pal5, aumentou o peso fresco da parte aérea da planta em 28% [Baldani et al., 1999]. Outra evidência da contribuição de *G. diazotrophicus* na FBN foi demonstrada pela inoculação de plantas de cana-de-açúcar micropropagadas com a estirpe selvagem Pal5 e seu mutante não fixador (*nif-*), seguida da medição de ^{15}N incorporado aos tecidos vegetais [Sevilla et al., 2001]. Oliveira e colaboradores demonstraram que a inoculação combinada de bactérias diazotróficas endofíticas e associativas promove um efeito sinérgico quando comparado com a inoculação individual da bactéria em plantas micropropagadas de cana-de-açúcar [Oliveira et al., 2002]. Aumentos da ordem de 30% no acúmulo de nitrogênio via FBN foram observados nestas plantas. Além da variedade existente entre estirpes de bactérias na eficiência durante a associação, existem também diferenças entre os diversos cultivares de cana-de-açúcar na capacidade de obtenção do nitrogênio via FBN. Recentemente, Oliveira e colaboradores demonstraram, em cana-de-açúcar, que a contribuição da FBN foi influenciada pela combinação das espécies presentes no inóculo, genótipo da planta, tipo de solo e fertilização de nitrogênio. As diferenças de FBN entre os diferentes cultivares indicam que fatores genéticos da planta regulam a eficiência do processo [Oliveira et al., 2006].

Desse modo, essa bactéria torna-se um modelo de estudo de organismos fixadores de nitrogênio, tanto pela compreensão da fixação biológica de nitrogênio (FBN) propriamente dita quanto pelo seu potencial biotecnológico como, por exemplo, pela possibilidade de substituição de fertilizantes nitrogenados em diferentes culturas de importância econômica, principalmente na cana-de-açúcar. De fato, tem sido largamente documentado que *G. diazotrophicus* pode aumentar a taxa de crescimento da cana-de-açúcar, e este benefício é associado a algumas características fisiológicas desta espécie:

- A primeira delas é a capacidade de fixar nitrogênio atmosférico. A transferência

do nitrogênio bacteriologicamente fixado ao vegetal contribui com uma parcela do nitrogênio necessário para o desenvolvimento da planta hospedeira,

- A produção de fatores estimulantes do crescimento vegetal, destacando-se a produção dos fitormônios auxina (como o ácido 3-indol acético, IAA) e giberilina, que têm sido amplamente detectados em culturas deste endófito [Lee et al., 2000, Muñoz-Rojas e Caballero-Mellado, 2003, Muthukumarasamy et al., 2006].
- Aumento da disponibilidade de nutrientes através da solubilização de Fósforo (P) e Zinco (Zn) [Linu et al., 2009, Saravanan et al., 2007].
- Outro exemplo de importância biotecnológica está relacionado a processos de controle biológico, devido a uma atividade antagonista contra *Xanthomonas albilineans*, como resultado da produção de uma bacteriocina similar à lisozima. A *G. diazotrophicus* é capaz de inibir o crescimento deste fitopatógeno responsável pela doença da escaldadeira folia na cana-de-açúcar, possivelmente através da destruição da parede celular da cepa patogênica [Piñón et al., 2002, Blanco et al., 2005].
- Além disso, a colonização por *G. diazotrophicus* leva a um aumento da densidade de pêlos radiculares, da taxa de raízes secundárias e da superfície radicular [Bashan e Levanony, 1990] que resulta num aumento da absorção de água e nutrientes, aumentando assim a capacidade da planta de produzir e suportar estresses ambientais [Baldani et al., 1983, Kapulnik et al., 1985].
- Outro aspecto que deve ser destacado também é que, diferentemente da interação rizóbio/ leguminosa, na associação *G. diazotrophicus*-planta não há formação de nódulos ou qualquer outra estrutura formada ou induzida pela bactéria, o que sugere uma interação muito bem sucedida no interior da planta promovendo o crescimento vegetal sem causar nenhum sintoma de doença, cujo mecanismo molecular precisa ser melhor entendido [Baldani e Baldani, 2005].

1.3.2 Isolamento e habitat natural

O primeiro isolamento do microrganismo utilizado neste estudo foi feito por Cavalcante e Döbereiner, a partir de raízes e caules de diversas variedades de cana-de-açúcar-de-açúcar (*Saccharum officinarum*) cultivadas em diferentes regiões do Brasil. O organismo foi descrito como “uma nova bactéria ácido-tolerante fixadora de nitrogênio” e foi nomeada *Saccharobacter nitrocaptans* [Cavalcante e Döbereiner, 1988]. Após seu primeiro isolamento, a associação de *G. diazotrophicus* com a cana-de-açúcar foi confirmada por novos isolamentos a partir de variedades de cana-de-açúcar cultivadas comercialmente na Austrália [Li e MacRae, 1991], México [Fuentes-Ramirez et al., 1993] e de plantas provenientes de Cuba cultivadas em estufas no Canadá [Dong et al., 1995]. Além da cana-de-açúcar, *G. diazotrophicus* foi isolada de outras plantas que são propagadas vegetativamente e alta sacarose como por exemplo o capim Cameroon (*Pennisetum purpureum*) [Döbereiner et al., 1993], batata (*Ipomoea batatas*) [Paula et al., 1991, Paula et al., 1992], sorgo (*Sorghum vulgare*) [Paula et al., 1991] e abacaxi (*Annanas sp.*) [Tapia-Hernández et al., 2000]. Até 1991, a ausência *G. diazotrophicus*, na rizosfera de cana-de-açúcar [Cavalcante e Döbereiner, 1988], assim como sua ausência no interior de plantas diferentes da cana-de-açúcar cultivadas junto com a cana-de-açúcar [Reis et al., 1994] reforçaram a idéia da do caráter endofítico desta bactéria e de sua associação exclusiva com plantas que acumulam sacarose e que propagam vegetativamente [Döbereiner et al., 1993]. No entanto, a partir de 1991 a *G. diazotrophicus* foi detectada ocasionalmente na rizosfera de cana-de-açúcar [Li e MacRae, 1991], na rizosfera e tecidos de plantas de café (*Coffea arabica L.*) [Jiménez-Salgado et al., 1997] e do cereal *Eleusine coracana* [Loganathan et al., 1999], que não contém alta concentração de sacarose e sua propagação é por sementes, bem como *Saccharococcus sacchari*, um inseto-praga da cana-de-açúcar [Ashbolt e Inkerman, 1990, Caballero-Mellado et al., 1995].

1.3.3 Taxonomia

Ao primeiro isolamento, o microrganismo utilizado neste estudo foi nomeado *Saccharobacter nitrocaptans* [Cavalcante e Döbereiner, 1988], mas com base em experimentos de hibridização RNA/DNA e DNA/DNA seu nome foi modificado para *Acetobacter nitrocaptans*. Foi só alguns meses depois que Gillis e colaboradores, baseados em estudos de hibridização de DNA, na presença de flagelação lateral e na capacidade de oxidar o acetato e lactato a CO_2 e água, sugeriram o nome oficial de *Acetobacter diazotrophicus*, nome com o qual ainda hoje continua a se reconhecer esta espécie, e definir o tipo de estirpe LMG 7603 (Döbereiner PAL 5^T = ATCC 49037) [Gillis et al., 1989]. Tradicionalmente a família *Acetobacteraceae* era dividida nos gêneros *Acetobacter* e *Gluconobacter* [Swings, 1992]. No entanto, em 1997, Yamada e colaboradores propuseram um novo regime na taxonomia da família, e a dividiram nos gêneros: *Acetobacter*, *Acetobacteraceae*, *Gluconobacter*, *Gluconoacetobacter* e *Acidomonas*, com base na sequência parcial do gene 16S do RNA ribossomal e do tipo ubiquinona predominante [Yamada et al., 1997]. Em 1998, Yamada e colaboradores corrigiram o nome do gênero *Gluconoacetobacter* para *Gluconacetobacter*, e foi só a partir de então que a primeira estirpe fixadora de nitrogênio descrita nesta família passou a ser conhecida oficialmente como *Gluconacetobacter diazotrophicus* [Yamada et al., 1998].

A família *Acetobacteraceae* pertence a uma subclasse α -Proteobacteria [Young, 1992]. São bactérias aeróbicas, Gram negativas, caracterizadas por sua capacidade de oxidar o etanol a ácido acético em meio de cultura com pH neutro ou ácido [De Ley et al., 1984, Swings, 1992]. Genotipicamente a família *Acetobacteraceae* pode ser diferenciada de outras α -Proteobacterias pela presença de dois sítios de restrição SphI e NcoI dentro do gene 16S do DNA ribossomal [Caballero-Mellado et al., 1999, Jiménez-Salgado et al., 1997]. Os gêneros *Acetobacter*, *Gluconacetobacter* e *Acidomonas*, ao contrário *Gluconobacter*, são capazes oxidar completamente etanol e lactato a CO_2 e H_2O (Swings, 1992). Embora o gênero *Gluconacetobacter* compartilhe com outros membros de sua família a capacidade de oxidar etanol, ele é o único gênero da família (descrito até agora), que tem espécies com

a capacidade de fixar nitrogênio atmosférico [Swings, 1992, Fuentes-Ramírez et al., 2001, Gillis et al., 1989].

1.3.4 Características fisiológicas

Gillis e colaboradores descreveram a espécie *G. diazotrophicus* como uma bactéria Gram negativa, aeróbica, com formato de bastonetes com extremidades arredondadas, com tamanho variando entre 0,7 e 0,9 μM , móveis e com flagelos laterais ou peritríquios [Gillis et al., 1989]. Seu crescimento é ótimo a 30°C, pH 5,5 e 10% de sacarose (correspondente à concentração de sacarose em que este organismo está exposto em seu habitat principal, a cana-de-açúcar), mas tolera até 30% deste açúcar [Cavalcante e Döbereiner, 1988, Li e MacRae, 1991].

Por poder reduzir o nitrogênio atmosférico ela pode crescer em meio sem aminoácidos ou sais de amônio. Ela pode crescer e fixar N_2 em valores de pH entre 2,5 e 7,0 [Stephan et al., 1991]. A atividade de nitrogenase é apenas parcialmente inibida pela adição de 2,65 g/L de $(\text{NH}_4)_2\text{SO}_4$ e vários aminoácidos mostram um efeito similar na presença de 5,0 g/L de glicose [Stephan et al., 1991]. Em níveis de 0,26, 0,66 e 1,32 g/L $(\text{NH}_4)_2\text{SO}_4$ na presença de 1% de sacarose, a atividade da nitrogenase é parcialmente inibida, mas as células mantem a atividade de nitrogenase quando a concentração de sacarose é de 10% [Reis e Döbereiner, 1998]. Esta proteção poderia ser explicada pela baixa assimilação de amônia quando as células de *G. diazotrophicus* são cultivadas com 10% de sacarose. Esse mecanismo é muito interessante para a bactéria, pois considerando-se que o seu habitat é geralmente rico em sacarose, sua nitrogenase pode estar ativa mesmo quando ela está em um ambiente rico em amônia [Reis e Döbereiner, 1998].

A *G. diazotrophicus* não utiliza nitratos (NO_3^-) devido à ausência de uma nitrato redutase. O processo de FBN não é afetado pela presença de 10 mM de NO_3^- [Stephan et al., 1991, Li e MacRae, 1991]. Ela também é catalase (+) e oxidase (-). Quando é crescida no meio LGI com azul de bromotimol forma colônias laranja escuro, e forma uma película de cor amarela no meio LGI semi-sólido, um aspecto que diferencia

esta espécie de outras bactérias diazotróficas [Alvarez e Drets, 1995]. Quando é crescida em meio com $CaCO_3$ forma um halo de solubilização [Cavalcante e Döbereiner, 1988]. A espécie *G. diazotrophicus* pode ser diferenciada de outras espécies fixadoras de nitrogênio pela formação de colônias de cor marrom escuro (após 5 a 7 dias de incubação), com bordas claras, quando crescidas em meio de cultura com extrato de batata [Fuentes-Ramírez et al., 2001].

Ela cresce bem em meios de cultura sintéticos contendo sacarose, glicose, frutose, gluconato, lactato, piruvato, acetato, manitol, sorbitol ou glicerol como única fonte de carbono [Gillis et al., 1989, Stephan et al., 1991, Ureta et al., 1995]. No entanto, não cresce com ácidos dicarboxílicos como única fonte de carbono [Alvarez e Drets, 1995, Stephan et al., 1991]. Acredita-se que a incapacidade de crescer em ácidos dicarboxílicos se deva à ausência de um sistema de transporte ou a perda de uma das enzimas da via que os catalizam já que Alvarez e colaboradores demonstraram a presença de algumas destas enzimas em extratos celulares de *G. diazotrophicus* [Alvarez e Drets, 1995]. Ela não cresce em dissacarídeos como lactose, melobiose, celobiose e maltose [Ureta et al., 1995].

A sacarose é a principal fonte de carbono que a bactéria tem em seu habitat principal, o apoplasto da cana-de-açúcar [Dong et al., 1994, Dong et al., 1995]. Dada a ausência de um sistema de transporte que traga este açúcar para o interior da célula [Alvarez e Drets, 1995], seu crescimento é baseado na ação de uma enzima extracelular que o hidrolisa: a levansucrase (sacarose 6- fructosiltransferase, EC. 2.4.1.10), que é expressa constitutivamente [Hernández et al., 1995, Hernández et al., 2000]. Esta enzima hidroliza a sacarose e transfere o resíduo frutósila para outra sacarose para formar frutanos intermediários que logo se polimerizam em compostos de alto peso molecular [Cote e Ahlgran, 1993]. No caso particular da glicose, Alvarez e Martínez-Drets [Alvarez e Drets, 1995] encontraram, em estudos com glicose radioativa, que a marca era eliminada do meio de cultura rapidamente, indicando a presença de um sistema de transporte da glicose para o interior da bactéria. No entanto, a conversão de glicose para o seu correspondente gluconato, sem a entrada prévia da mesma dentro da cé-

lula, parece ser a via mais relevante do metabolismo deste açúcar [Alvarez e Drets, 1995, Attwood et al., 1991].

Capítulo 2

Identificação e análise dos transportadores

2.1 Introdução

2.1.1 O sequenciamento de genomas

O desenvolvimento do equipamento e das técnicas que permitem o sequenciamento automatizado de fragmentos de DNA e conseqüentemente de genomas completos é a base do aparecimento da área científica da Genômica. Desde 1995, quando foi disponibilizada a primeira seqüência completa do genoma de um ser vivo (da bactéria *Haemophilus influenza*), as abordagens genômicas e pós-genômicas vêm revolucionando a forma de investigar e de desenvolver qualquer outra atividade nas várias vertentes da Biologia. De acordo com o banco de dados GOLD™ - Genomes OnLine Database v. 3.0 - existem hoje (março de 2010) cerca de 1100 genomas completos publicados, incluindo o genoma humano (Figura 2.1). Esse número ainda cresce de maneira exponencial e acredita-se que até o final de 2010, cerca de 5000 seqüências estejam disponíveis - Figura 2.2, [Liolios et al., 2009].

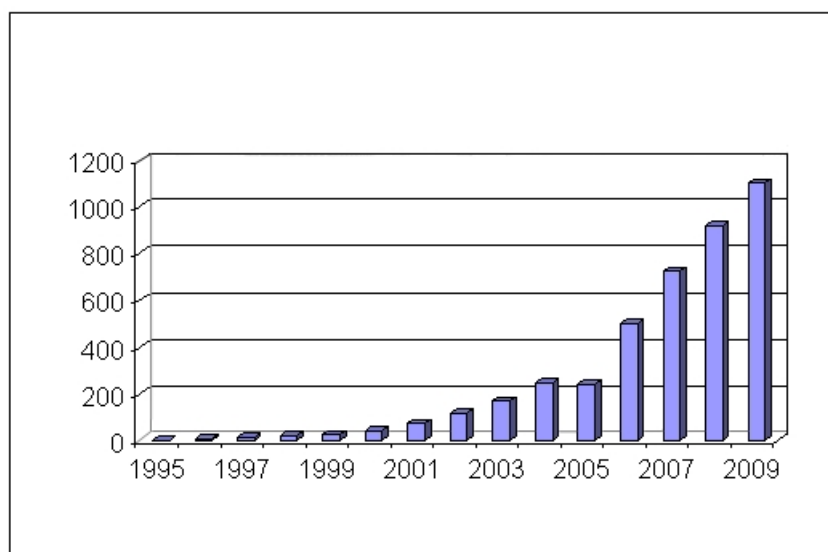


Figura 2.1: Número genomas completos sequenciados até setembro de 2009 (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)

O sequenciamento de genomas completos ganha cada vez mais importância no cenário científico pois fornece informações detalhadas sobre uma espécie que não podem facilmente ser gerada por meios alternativos. Estas informações incluem a presença de genes únicos, de proteínas hipotéticas conservadas, a composição de GC, o índice de variações totais e de elementos repetitivos e/ou de inserção, além de permitir uma reconstrução detalhada do perfil fisiológico e metabólico dos organismos estudados, e permitir uma análise funcional subsequente dos genes, dos quais nenhuma informação precedente estava disponível.

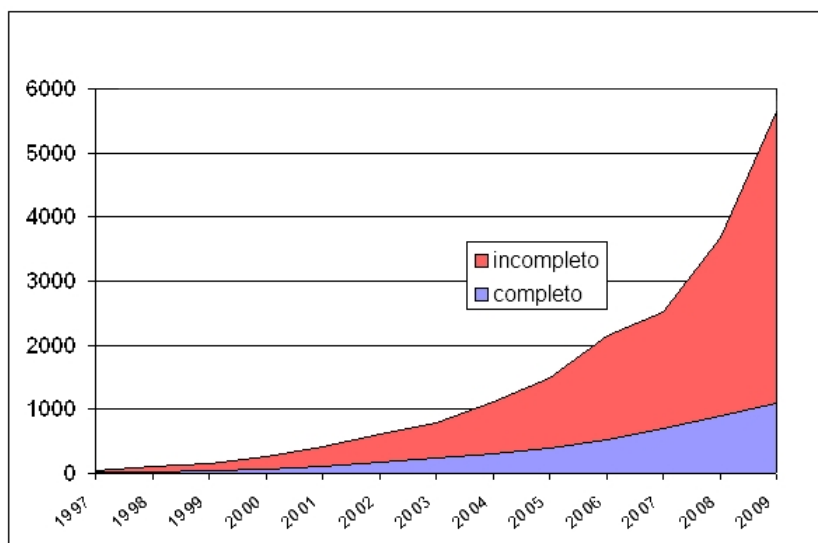


Figura 2.2: Número de projetos de sequenciamento de genoma depositados no GOLD™ (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)

O Brasil, desde a publicação do genoma da *Xylella fastidiosa* [Simpson et al., 2000], vem crescentemente aumentando a sua participação mundial em projetos de sequenciamento de genomas. Uma grande parte dos projetos de sequenciamento de genomas tem algum interesse biotecnológico (Figura 2.3) [Santos et al., 2005] e, devido ao fato de grande parte da economia brasileira ser baseada na atividade agropecuária, quase a totalidade dos diferentes projetos de sequenciamentos de genomas desenvolvidos em território nacional são voltados para esse setor da economia [Carraro e Kitajima, 2002]. Seguindo essa tendência nacional, em 2000, a Fundação Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro (FAPERJ) lançou o projeto de sequenciamento do genoma completo da bactéria *G. diazotrophicus*. Além do sequenciamento desta bactéria o projeto visava estabelecer uma infraestrutura de pesquisa em diversas instituições, criando uma rede de laboratórios com competência nas áreas de genômica, de biologia estrutural e de bioinformática.

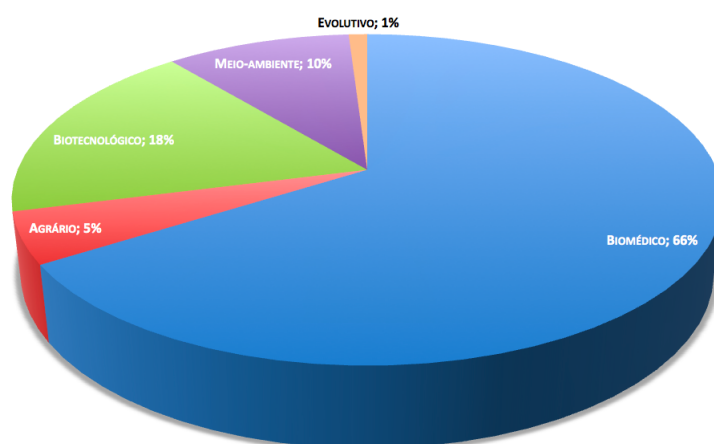


Figura 2.3: Principal relevância de projetos de genoma bacterianos realizados no mundo (modificada de http://www.genomesonline.org/gold_statistics.htm - acesso em março de 2010)

O consórcio de laboratórios RIOGENE ficou responsável pelo projeto e consistia em laboratórios do Departamento de Bioquímica da UFRJ, Departamento de Informática da PUC-RJ; Instituto de Biofísica Carlos Chagas Filho da UFRJ; Instituto de Biologia da UFRJ; Universidade Estadual do Norte Fluminense (UENF); Universidade Federal Rural do Rio de Janeiro (UFRRJ); Instituto de Biologia Roberto Alcântara Gomes (IBRAG/Uerj) e o Núcleo de Fixação Biológica de Nitrogênio do centro Nacional de Pesquisa em Agrobiologia (CNPAB) da Embrapa.

O processo de sequenciamento do genoma de *G. diazotrophicus* pode ser dividido em três etapas principais. Elas foram:

1. A quebra do DNA inteiro do organismo em fragmentos pequenos, através da técnica de “shotgun”, que foram clonados em vetores plasmidiais (pUC18) ou cosmídeos (pLARF3) que foram sequenciados em suas extremidades. Um total de 103506 sequências de alta qualidade foram obtidas e montadas em contigs utilizando-se o programa de montagem Phrap.
2. Para o fechamento de regiões não cobertas pelo sequenciamento aleatório, 16963 sequências adicionais foram obtidas pelo método de “primer walking” ou PCR direcionado. A edição manual das sequências foi realizada usando-se o programa GAP4

[Bonfield et al., 1995].

3. Uma vez obtidos os dados do sequenciamento das moléculas de DNA é preciso saber o que representa cada uma das sequências nucleotídicas produzidas. A anotação consiste no processo de identificação dessas sequências.

2.1.2 Anotação de genomas

O processo de anotação apresenta 4 etapas distintas [Schwartz, 2000][Koonin et al., 1996, Downs e Escalante-Semerena, 2000]:

- A primeira consiste na anotação em nível de nucleotídeos, cujo principal objetivo é a localização física de ORFs (fases de leitura aberta) nas sequências de DNA, assim como elementos repetitivos, promotores e sítios de ligação ribossomais. O principal objetivo dessa etapa é construir um mapa do genoma do organismo, posicionando cada um dos possíveis genes e caracterizando as regiões não-gênicas. Nesta fase, alguns programas de predição gênica são usados para a localização de possíveis genes nas sequências de DNA. A procura por elementos como o códon de iniciação de proteínas (a trinca de nucleotídeos ATG, por exemplo) e códons de terminação na mesma fase de leitura são utilizados por alguns desses programas, como o GLIMMER. O tamanho delimitado por esta janela de leitura é utilizado para definir uma determinada região como sendo um gene.
- A segunda etapa da anotação ocorre em nível de proteínas, cujos objetivos residem na determinação de identidades e prováveis funções dos genes, visando identificar genes presentes (e eventualmente ausentes) num determinado genoma. Essa etapa consiste em identificar quais proteínas são codificadas, e nisso consiste o processo de anotação das sequências proteicas. Nessa etapa, procura-se montar um catálogo dos genes presentes no organismo estudado, dando-lhes nomes e associando-os a prováveis funções. No caso do genoma de *G. diazotrophicus*, antes da anotação manual de cada gene, diferentes ferramentas foram usadas: A busca por sequências similares foi rea-

lizada em diferentes bancos de dados incluindo Uniprot [UniProt Consortium, 2008], PROSITE [Falquet et al., 2002], Pfam [Bateman et al., 2004] e Inter- Pro [78]. Além destes bancos, o SignalP [Emanuelsson et al., 2007], TMHMM [Krogh et al., 2010] e tRNAscan-SE [Lowe et al., 1997] também foram usados. Todos os dados foram visualizados utilizando-se o programa Artemis [Carver et al., 2008].

- A terceira etapa ocorre em nível de processo, onde os grupos de pesquisa buscam realizar a identificação de diversas vias metabólicas.
- A quarta etapa consiste na análise das sequências reguladoras presentes no genoma, visando implementar uma anotação funcional eficiente, e aprimorar a compreensão das diversas vias metabólicas presentes no organismo, em especial aquelas que apresentam importância econômica, científica, tecnológica e acadêmica.

2.1.3 Ferramentas de bioinformática

Dentre as principais ferramentas de bioinformática utilizadas para execução do processo de anotação, destacam-se:

- BLAST (Basic Local Alignment Sequence Tool), que consiste numa ferramenta utilizada para verificação de homologia de uma determinada sequência, com sequências estudadas previamente por outros pesquisadores [Altschul et al., 1998, Wu-blast, 2006];
- RBSFinder, que localiza sítios de ligação ribossomal [Tigr, 2006];
- programas de predição de genes como o GLIMMER [Delcher et al., 1999] e o ORF-Finder [NCBI, 2006a];
- algoritmos para predição de estruturas de proteínas como o Modeller [Sali, 2006] e o Threader [Jones, 2006];
- pacotes de análise filogenética como o PHYLIP (PHYLogeny Inference Package) [Felsenstein, 2006];

- programas projetados para comparação de vários genomas ex: Alfresco [Sanger, 2006].

Com relação aos bancos de dados, podemos destacar os que armazenam sequências de nucleotídeos, tais como o Genbank-NT [NCBI, 2006b], GSDB [Harger et al., 1998], GDB [Letovsky et al., 2006] e o EMBL [Coschrane et al., 2006]; os que armazenam sequências de aminoácidos como o SWISS-PROT [Boeckmann et al., 2003], PIR [Wu et al., 2002] e o Genbank-NR [NCBI, 2006b]; os bancos de dados de estruturas terciárias de proteínas como o PDB - Protein Data Bank [Berman et al, 2000]; padrões de sequências proteicas conservadas associadas com funções específicas num organismo, tais como o PROSITE [Falquet et al., 2002]; e bibliotecas que oferecem padrões de sequências de estruturas de domínios protéicos mais longos, como o PFAM [Bateman et al., 2004].

Em linhas gerais, podemos afirmar que a anotação de um genoma consiste num processo interativo, que requer a utilização de diversas ferramentas de bioinformática, com finalidade de refinar os dados brutos gerados a partir de um projeto de sequenciamento genômico [Bionotes, 2006], através da utilização de ferramentas que possibilitem a localização de regiões codificantes numa sequência de DNA [Delcher et al., 1999, NCBI, 2006a], a predição de função e estrutura da proteína codificada por um determinado gene [Jones, 2006, Sali, 2006], assim como a determinação de relações filogenéticas entre diferentes organismos [Felsenstein, 2006, Sanger, 2006]. Desta forma, são elucidadas questões relacionadas com organização genômica, aspectos bioquímicos, fisiológicos e evolutivos do organismo em estudo, e de espécies relacionadas e/ou espécies que ocorrem em associação [Koonin et al., 1996, Moszer, 1998].

2.1.4 Transportadores e sistemas de transporte

A vida envolve tanto interconversões químicas quanto fluxos através das membranas biológicas, que referimos como metabolismo e transporte e nenhuma descrição de vida é completa sem uma compreensão inclusa de ambos fenômenos [Saier, 2000]. Assim, não é estranho que todos os organismos apresentem cerca de um terço de suas proteínas como proteínas de membrana, e cerca de um terço destas são proteínas de transporte

[Werten et al., 2002]. Assim, os sistemas de transporte em membranas têm papéis fundamentais no metabolismo e atividade celular, por catalisarem a translocação de solutos usando distintos mecanismos moleculares, e funcionam na fisiologia celular em processos distintos, tais como a aquisição de nutrientes orgânicos, e a manutenção da homeostase iônica, a extrusão de compostos tóxicos, a sinalização e comunicação celular e ambiental, entre outros [Saier, 2000]. Desta maneira pode-se dizer que repertório de proteínas de transporte da membrana determina o estilo de vida de uma espécie [Ren e Paulsen, 2005].

Os vários sistemas de transporte existentes diferem em relação a sua topologia na estrutura das membranas biológicas, aos mecanismos de acoplamento energético e transporte, e a especificidade por substrato [Paulsen et al., 1998]. Contudo, análises filogenéticas entre os diferentes grupos de transporte demonstram que uma classe é, frequentemente, mais homóloga a uma proteína da mesma classe, em outro organismo distante evolutivamente, que a outra classe de transportador no mesmo organismo. Então, parece que pelo menos a especificidade de acoplamento de energia e por substrato evoluiu muito cedo, e antes da divisão dos três domínios da vida: Bacteria, Archaea e Eucarya [Palmgren e Axelsen, 1998].

Isso permite que os transportadores sejam classificados pelo TCDB (Transporter Classification System; <http://www.tcdb.org>) em famílias e subfamílias, de acordo com sua função, filogenia e/ou especificidade de substrato. Os mecanismos de acoplamento energético e transporte servem como base preliminar para a classificação, por causa de suas características relativamente estáveis [Saier, 2000].

O website do TCDB detalha um sistema de classificação global de proteínas de transporte de membrana conhecido como a sistema de classificação do transportadores (TC, de “Transport Classification”). O sistema TC é análogo ao sistema adotado para a classificação de enzimas (EC, de “Enzyme Commission”) no entanto ele incorpora tanto informação funcional quanto filogenética. Descrições, números de TC, e exemplos das mais de 360 famílias de transportadores são fornecidas. A classificação dos transportadores de *G. diazotrophicus* foi realizada seguindo o modelo TCDB que apresenta cinco classes

principais de transportadores [Saier et al., 2006]. Elas são:

1. **Canais/ Poros** - Os canais são proteínas complexas que possuem sítios de translocação (poros que atravessam as membranas) para íons específicos, os quais podem ser abertos ou fechados por alterações na conformação da proteína. Os canais possibilitam a passagem de um grande número de íons, sempre a favor de um gradiente de potencial eletroquímico (diferenças de concentração de íons e de potencial elétrico), e funcionam como um portão. O tamanho do poro e a densidade das cargas na superfície interior do canal determinam a afinidade do íon a ser transportado, em uma região do canal denominada filtro de seletividade. Um único canal aberto pode permitir a passagem de mais de 100.000.000 de íons (10^8 íons) por segundo. Entretanto, os "portões" abrem ou fecham o canal em resposta a estímulos externos, e o tempo de abertura de um canal ser muito curto.
2. **Transportadores ativos secundários** ("Electrochemical Potential-driven transporters") - São proteínas carreadoras que usam o gradiente de de íons como energia. O transporte ativo secundário usa a energia liberada pela difusão de um íon para conseguir o movimento contra gradiente de outro soluto pela membrana. Para a energia da difusão iônica ser aproveitada para o transporte ativo, ambos os processos tem que ser mediados pela mesma proteína de membrana. São estereoespecíficos.
3. **Transportadores ativos primários** - Proteínas carreadoras que usam ATP como energia. Também são comumente chamadas de "bombas". Elas são, ao mesmo tempo, proteínas carreadoras e ATPases. As bombas são proteínas integrais ativadas por energia química (ATP) ou luminosa. Estas bombas também são chamadas de bombas eletrogênicas, por gerarem diferenças de potencial elétrico através das membranas, sendo responsáveis pela geração de um gradiente de prótons através das membranas (transporte ativo primário).
4. **Translocadores** - Translocação é uma forma de transporte ativo que ocorre em procariotos. Na translocação a substância transportada torna-se quimicamente al-

terada durante o seu transporte através de uma membrana de modo que uma vez lá dentro, a membrana citoplasmática torna-se impermeável a essa substância e ela permanece dentro da célula. Um exemplo de translocação em bactérias é o sistema fosfotransferase. Um grupo fosfato de alta energia é transferido a partir de fosfoenolpiruvato por uma série de enzimas até a glicose. A enzima final fosforila a glicose e transporta-a através da membrana como glicose-6-fosfato.

5. **Carreadores de elétrons** - Catalizam o fluxo de elétrons através da membrana a partir de doadores localizados em um lado da membrana para receptores localizados no outro lado da membrana. Dependendo da direção do fluxo de elétrons esse sistema pode aumentar ou diminuir o potencial da membrana, sendo desta maneira muito importantes para o metabolismo energético celular.
6. NÃO UTILIZADO. Os números 6 e 7 não são utilizados para representar nenhuma classe do TCDB
7. NÃO UTILIZADO.
8. **Auxiliadores de transporte** - Proteínas que, de alguma forma facilitam o transporte através de membranas biológicas, mas não participam diretamente do transporte.
9. **Sistemas de transporte ainda não caracterizados** - Classe distinta, onde são incluídos transportadores cujo mecanismo energético ou de função ainda são desconhecidos.

Apesar das proteínas de transportes serem vitais à compreensão da potencialidade metabólica de organismos, poucos grupos têm se dedicado à análise sistemática dos sistemas de transportes dos microorganismos, em especial daqueles em que os genomas completos foram publicados e/ou ainda uma análise comparativa dos sistemas de transportes entre os mais diferentes microorganismos [Ren e Paulsen, 2005].

Tendo em vista uma análise mais detalhada do genoma da *G. diazotrophicus*, a condução de um inventário de todas as proteínas de transporte é uma etapa chave para entender as estratégias desenvolvidas por este microorganismo para adaptar-se às circunstâncias ambientais, com interesse adicional para o estudo das proteínas de transporte que estejam envolvidas no processo de associação com o vegetal.

Durante o processo de anotação do genoma da *G. diazotrophicus* PAL5 os diferentes grupos que integram o RioGene ficaram responsáveis pela identificação de genes de sua proficiência e interesse. Desse modo, nosso grupo ficou responsável pela identificação de ORFs relacionadas com os sistemas de transporte em membranas.

2.2 Objetivos

Essa parte da tese tem como objetivos:

- Identificação de ORFs relacionadas com o sistema de transportadores em membrana, durante o processo de anotação do genoma de *G. diazotrophicus*.
- Análise dos grupos de transportadores identificados.
- Análise comparativa entre a frequência dos transportadores de *G. diazotrophicus* e de outras bactérias já sequenciadas.

2.3 Resultados e discussão

A mineração dos resultados da anotação de todas as ORFs validadas, foi realizada através do acesso direto no banco de dados do sistema de anotação (Artemis), onde todas as informações pertinentes as ORFs validadas pelos membros integrantes do Projeto RioGene foram acessadas para desenvolvimento de análises estatísticas e observação das principais particularidades do genoma deste importante endófito (detalhes na Sessão 5.1.1).

Vale destacar como características gerais do genoma do *G. diazotrophicus* PAL5 (GDI) que seu genoma completo é composto por um cromossomo circular de 3.944.163 pares de base (pb), com um conteúdo GC médio de 66,19%, e dois plasmídeos de 38.818 e 16.610 pb, respectivamente. O cromossomo circular tem um total de 3.864 sequências codificantes previstas (CDS), com uma capacidade global de codificação de 90,67%. Entre as ORFs identificadas foram atribuídas uma função provável (putativa) a 2.861 delas, e 1.077 codificam proteínas hipotéticas. Em relação ao RNA não-codificante, 12 rRNAs (quatro operons do rRNA) e 55 tRNAs foram identificados. O plasmídeo maior (pGD01) tem 53 CDS; aproximadamente 70% codificam proteínas hipotéticas ou hipotéticas conservadas e cinco codificam proteínas relacionadas a funções do plasmídeo. As 11 CDS restantes codificam para os componentes de um provável sistema de secreção do tipo IV (T4SS). O plasmídeo pequeno (pGD02) tem 21 CDS, e cerca de 50% delas são proteínas hipotéticas [Bertalan et al, 2009].

A análise das 1160 proteínas previstas como proteínas de membrana mostrou que 563 delas eram relacionadas ao sistema de transporte. Elas são mostradas na Tabela 2.1.

Tabela 2.1: Descrição das ORFs identificadas com pertencentes ou relacionadas com o sistema de transportes no genoma da *G. diazotrophicus*

ORFs	Função da seqüência	TC Number
GD0019	putative sugar transporter protein	2.A.1.1
GD0022	putative cation transporter protein	2.A.4
GD0042	putative sugar transporter protein	2.A.1
GD0046	Outer membrane protein (Omp A domain)	1.B.17
GD0056	putative DNA uptake protein	9.A.21
GD0058	putative ABC transporter	3.A.1.210
GD0096	putative membrane protein	9.B
GD0105	putative membrane protein	9.B
GD0131	traG Conjugal transfer protein TRAG	3.A.7
GD0133	trbB type II secretion system protein	3.A.5
GD0134	trbB conjugal transfer protein trbB	3.A.7
GD0135	trbE conjugal transfer protein trbE	3.A.7
GD0136	trbJ putative type IV secretion system protein	3.A.7
GD0137	trbL putative type IV secretion system protein	3.A.7
GD0138	trbF putative type IV secretion system protein	3.A.7
GD0139	trbG putative type IV secretion system protein	3.A.7
GD0140	trbI putative type IV secretion system protein	3.A.7
GD0146	mdtB Multidrug resistance protein	2.A.6.2
GD0147	mdtA Multidrug resistance protein	2.A.6.2
GD0148	RND efflux system, outer membrane lipoprotein, NodT	2.A.6.2
GD0169	putative membrane protein	9.B
GD0195	putative MFS permease	2.A.1.14
GD0198	TonB-dependent receptor	1.B.14
GD0199	putative sugar transporter	2.A.1.6
GD0210	accB Biotin carboxyl carrier protein of acetyl-CoA	3.B.1
GD0214	putative glucarate transporter	2.A.1.14
GD0215	TonB-dependent receptor	1.B.14
GD0222	NAD(P)H dehydrogenase (quinone)	2.A.37
GD0232	nhaA Na(+)/H(+) antiporter 1	2.A.33
GD0243	Multidrug resistance protein	2.A.1.2
GD0244	secretion protein HlyD	2.A.6
GD0245	outer membrane protein	2.A.6.2
GD0257	putative ABC-type transporter, substrate binding protein	3.A.1.5
GD0259	Inner membrane ABC transporter permease protein	3.A.1.5
GD0260	Dipeptide transport system permease protein	3.A.1.5
GD0261	ABC transporter ATP-binding protein	3.A.1.5
GD0265	TonB-dependent receptor	1.B.14
GD0276	ptsP Phosphoenolpyruvate-protein phosphotransferase ptsP	8.A.7
GD0278	glpF Glycerol uptake facilitator protein	1.A.8
GD0291	Putative Na(+)/H(+) exchanger	2.A.36
GD0316	putative membrane protein	9.B
GD0331	putative transport protein	8.A
GD0332	putative transport protein	9.B
GD0343	corB putative Magnesium-transporting ATPase	3.A.3.4
GD0344	putative transport protein	1.A.30.2
GD0350	major facilitator superfamily transporter	2.A.1
GD0351	arsB Arsenical pump membrane protein	2.A.45
GD0352	major facilitator superfamily transporter	2.A.1.14

GD0357	pntA NAD(P) transhydrogenase subunit alpha part 1	3.D.2
GD0358	ptnB NAD(P) transhydrogenase subunit alpha part 2	3.D.2
GD0359	ptnB NAD(P) transhydrogenase subunit beta	3.D.2
GD0362	citA Citrate-proton symporter	2.A.1.6
GD0365	putative drug resistance transporter	2.A.1.3
GD0370	putative periplasmic binding proteins	3.A.1
GD0371	putative ABC transporter ATP-binding protein	3.A.1
GD0374	putative ribose transport system permease protein	2.A.7
GD0432	putative TonB-dependent receptor	1.B.14
GD0438	putative molybdenum import ATP-binding protein modC	3.A.1.8.1
GD0439	putative molybdenum transport system permease protein modB	3.A.1.8.1
GD0440	putative molybdate-binding periplasmic protein precursor	3.A.1.8.1
GD0470	Chemoreceptor mcpA	9.B
GD0471	Oxidoreductase (aldo/keto reductase) protein	8.A.5
GD0477	putative extracellular solute-binding protein	3.A.1
GD0479	putative ABC transporter, permease protein	3.A.1
GD0480	putative ABC transporter, permease protein	3.A.1
GD0481	putative ABC transporter, ATP-binding protein	3.A.1
GD0494	pseudopilin G precursor	3.A.15
GD0495	putative prepilin peptidase and N-methyltransferase O	3.A.15
GD0496	putative type II secretion pathway kinase GspE	3.A.15
GD0497	putative general secretion pathway protein F	3.A.15
GD0498	putative pseudopilin H precursor	3.A.15
GD0499	putative pseudopilin I precursor	3.A.15
GD0500	putative pseudopilin J precursor	3.A.15
GD0501	putative general secretion pathway protein L	3.A.15
GD0502	putative general secretion pathway protein M	3.A.15
GD0503	putative general secretion pathway protein N	3.A.15
GD0504	putative general secretion pathway protein D	3.A.15
GD0505	putative general secretion pathway protein K	3.A.15
GD0514	putative membrane protein	9.B
GD0517	putative membrane protein	9.B
GD0518	putative outer membrane protein (OmpA/MotB)	1.B.17
GD0520	putative MotA/ToIQ/ExbB proton channel family protein	1.A.30
GD0541	putative voltage-gated ClC-type chloride channel	1.A.11
GD0549	putative MFS transporter	2.A.1.2
GD0568	TonB-dependent receptor	1.B.14.1
GD0569	putative TonB-dependent receptor	1.B.14.1
GD0574	putative porin protein	1.B.5.1
GD0575	putative sulfate-binding protein precursor	3.A.1.6.1
GD0576	putative sulfate transport system permease protein	3.A.1.6.1
GD0577	putative sulfate transport system permease protein cysW	3.A.1.6.1
GD0578	putative sulfate/thiosulfate import ATP-binding protein cysA	3.A.1.6.1
GD0579	putative acriflavin resistance protein	2.A.6.2
GD0580	putative multidrug resistance protein mdtA precursor	2.A.6.2
GD0581	putative toluene efflux pump outer membrane protein ttgC precursor	2.A.6.2
GD0611	putative NADH-quinone oxidoreductase chain M	3.D.1
GD0622	putative lysine exporter protein	2.A.75.1
GD0624	putative sugar transporter	2.A.1.7
GD0627	putative galactose-proton symporter	2.A.1.1
GD0645	putative transporter protein	2.A.1.14
GD0652	putative membrane protein	9.B
GD0653	putative membrane protein	9.B

GD0656	kup1 potassium transport system protein	2.A.72
GD0666	galP Galactose-proton symporter	2.A.1.1
GD0667	conserved hypothetical protein	9.B
GD0670	mdtC Multidrug resistance protein	2.A.6.2
GD0671	hypothetical protein	9.B
GD0674	Iron permease protein	9.A.10
GD0678	mntH manganese transport protein mntH	2.A.55
GD0685	conserved hypothetical protein	9.B.63
GD0689	fecE Iron(III) dicitrate transport ATP-binding protein	3.A.1.14
GD0690	putative ferrichrome ABC transporter permease protein	3.A.1.14
GD0692	TonB-dependent receptor	1.B.14
GD0706	gudP putative glucarate transporter	2.A.1.14
GD0709	major facilitator superfamily (MFS) transporter	2.A.1
GD0714	Preprotein translocase secE subunit	3.A.5
GD0715	putative MFS transporter	2.A.1.3
GD0718	atpH ATP synthase delta chain	3.A.2.1
GD0719	atpA ATP synthase subunit alpha	3.A.2.1
GD0720	atpG ATP synthase gamma chain	3.A.2.1
GD0721	atpD ATP synthase subunit beta	3.A.2.1
GD0722	atpC ATP synthase epsilon chain	3.A.2.1
GD0731	amtB Ammonium transporter	2.A.49
GD0745	hypothetical protein	9.B
GD0761	major facilitator superfamily transporter	2.A.1
GD0790	Secretion protein HlyD	2.A.6
GD0791	putative membrane protein	2.A.6
GD0796	Porin B precursor	1.B.19
GD0803	putative surface antigen	9.B
GD0810	Purine nucleoside permease	2.A.41
GD0813	putative ABC transporter ATP-binding protein	3.A.1
GD0814	putative membrane protein	9.B
GD0821	conserved hypothetical protein	9.B
GD0825	secA Preprotein translocase secA subunit	3.A.5
GD0831	tolQ Protein tolQ	1.A.30
GD0832	tolR Biopolymer transport protein	1.A.30
GD0834	tolB Protein tolB precursor	1.A.30
GD0835	omp Outer membrane lipoprotein omp16	1.B.6.1
GD0838	ftsH Cell division protein ftsH	3.A.16
GD0846	putative membrane protein	9.B
GD0853	putative membrane protein	2.A.7
GD0857	tppB Tripeptide permease tppB	2.A.17
GD0901	putative aquaporin	1.A.8
GD0927	kefB Glutathione-regulated potassium-efflux system protein	2.A.37
GD0934	Porin B precursor	1.B.19
GD0945	hypothetical protein	5.A
GD0950	putative Hemolysin-type calcium-binding	1.C
GD1020	major facilitator transporter superfamily	2.A.1.2
GD1027	Major facilitator superfamily	2.A.1
GD1032	hlyD secretion protein HlyD	2.A.6
GD1034	multidrug efflux transporter protein	2.A.1.3
GD1040	Putative NAD(P)H oxidoreductase	2.A.37
GD1046	trbI conjugation TrbI-like protein	3.A.7
GD1047	trbG conjugal transfer protein trbG	3.A.7
GD1048	trbF conjugal transfer protein; TrbF	3.A.7

GD1049	conjugal transfer protein; TrbL	3.A.7
GD1051	trbJ conjugal transfer protein; TrbJ	3.A.7
GD1052	trbE conjugal transfer protein trbE	3.A.7
GD1053	trbB conjugal transfer protein; TrbB	3.A.7
GD1054	trbC Conjugal transfer protein TrbC	3.A.7
GD1055	Type II/IV secretion system protein	3.A.7
GD1057	putative membrane protein	9.B
GD1059	traG Conjugal transfer protein traG	3.A.7
GD1077	putative membrane protein	9.B
GD1084	pucl Probable allantoin permease	2.A.39
GD1139	putative ABC transport	3.A.1.12
GD1164	mdtC putative Multidrug resistance protein	2.A.6.2
GD1165	putative Purine-cytosine permease	2.A.39
GD1176	putative bicyclomycin resistance protein	2.A.1.2
GD1183	putative mechanosensitive ion channel	1.A.23
GD1193	putative sodium cotransporter	2.A.21.5.10
GD1199	putative MFS transporter	2.A.1
GD1200	probable TonB-dependent receptor	1.B.14.1
GD1213	Hypothetical transmembrane protein	9.B
GD1224	putative phosphoenolpyruvate-protein phosphotransferase	8.A.7.1
GD1225	putative phosphoryl transfer system, HPr	4.A.6
GD1226	Putative phosphotransferase enzyme IIA	4.A.6
GD1231	putative ABC transporter protein	3.A.1.105.2
GD1232	putative ABC transporter protein	3.A.1.105.2
GD1238	putative TonB-dependent Receptor protein	1.B.14.1
GD1238	putative sugar transport protein	2.A.1.6
GD1246	putative ABC transporter protein	3.A.1.15
GD1247	putative ABC transporter protein	3.A.1.15
GD1248	putative ABC transporter protein	3.A.1.15
GD1256	Putative ABC transporter permeases	3.A.1
GD1257	putative Magnesium transport protein corA	1.A.35.1
GD1257	putative ABC transporter, ATP-binding protein	3.A.1.6.1
GD1262	putative TonB-dependent receptor	1.B.14.1
GD1267	putative TonB-dependent Receptor protein	1.B.14
GD1276	putative Outer membrane protein	1.B.39.1
GD1283	putative membrane protein	2.A.66.1
GD1285	putative acriflavine resistance protein	8.A.1.6
GD1286	putative acriflavine resistance protein B	2.A.2.6
GD1326	putative Outer membrane protein oprM	2.A.6
GD1338	putative DNA translocase ftsK	3.A.12.1.2
GD1349	putative major facilitator protein	2.A.1
GD1356	putative sodium symporter protein	2.A.21
GD1380	putative tonB-dependent receptor	1.B.14.1
GD1385	putative antibiotic resistance transport membrane protein	2.A.1.2
GD1390	putative Biotin carboxyl carrier protein of acetyl-CoA carboxylase	3.B.1.1
GD1395	putative Heme exporter protein B	3.A.1.107.1
GD1396	putative cytochrome c biogenesis ATP-binding export protein ccmA	3.A.1.107
GD1398	putative ABC transporter protein	3.A.1
GD1399	putative amino acid ABC transporter, periplasmic amino acid-binding protein	3.A.1
GD1400	putative inner membrane amino-acid ABC transporter permease protein	3.A.1
GD1427	putative Biopolymer transport exbB protein	1.A.30
GD1428	putative Biopolymer transport	1.A.30
GD1429	putative Biopolymer transport exbD2 protein	1.A.30

GD1472	putative permease protein	2.A.40
GD1478	putative TonB-dependent receptor	1.B.14.1
GD1514	putative TonB-dependent Receptor protein	1.B.14.1
GD1529	D-ribose-binding protein	3.A.1.2.1
GD1536	Ribose transport system permease protein rbsC	3.A.1.2.1
GD1537	Ribose transport system permease protein rbsC	3.A.1.2.1
GD1538	Ribose transport ATP-binding protein rbsA	3.A.1.2.9
GD1539	Ribose ABC transporter, substrate-binding protein	3.A.1.2.1
GD1553	Copper resistance protein A	9.A.10
GD1554	Copper resistance protein B	9.A.10
GD1555	putative inner membrane transport protein	2.A.1.2
GD1558	putative heavy-metal ion transporter	2.A.6.1
GD1559	putative resistance protein	2.A.6
GD1560	putative type I secretion outer membrane protein,...	1.B.17
GD1572	putative hexuronate transporter	2.A.1.14
GD1574	Putative ABC transporter permease protein	3.A.1.2
GD1606	Probable potassium transport system protein kup	2.A.72
GD1608	Potassium-transporting ATPase A chain	3.A.3.7.1
GD1609	Potassium-transporting ATPase B chain	3.A.3.7.1
GD1610	Potassium-transporting ATPase C chain	3.A.3.7.1
GD1611	Sensor protein kdpD	2.A.1
GD1643	putative membrane protein	3.A.1
GD1644	putative extracellular solute-binding protein	3.A.1
GD1645	putative ABC-transporter ATP-binding protein	3.A.1
GD1652	Outer membrane protein	2.A.6.2.23
GD1660	outer membrane protein	2.A.6.2
GD1661	putative secretion protein	2.A.6
GD1662	putative transporter protein	2.A.1
GD1664	putative carbohydrate-selective porin	1.B.19.1.1
GD1687	putative oxidoreductase	8.A.5
GD1761	Type II/IV secretion system protein	3.A.7
GD1769	putative membrane protein	2.A
GD1770	putative membrane protein	2.A
GD1783	mscL Large-conductance mechanosensitive channel	1.A.22.1
GD1787	secD Protein-export membrane protein secD	2.A.6
GD1791	uup ABC transporter ATP-binding protein uup	3.A.1.120
GD1798	cydD Transport ATP-binding protein cydD	3.A.1.129
GD1799	cydC Transport ATP-binding protein cydC	3.A.1.129
GD1804	glpF Glycerol uptake facilitator protein	1.A.8
GD1814	tolC Outer membrane protein tolC precursor	1.B.17
GD1817	Putative nitrate transport ATP-binding protein	3.A.1.16
GD1818	putative glycine betaine transport system permease	3.A.1.12
GD1838	putative Na ⁺ /H ⁺ antiporter	2.A.36
GD1843	Mechanosensitive ion channel	1.A.23
GD1844	cationic amino acid transporter	2.A.3.3
GD1868	cvrA Cell volume regulation protein A	2.A.36
GD1895	putative transporter protein	2.A.28
GD1896	Magnesium and cobalt efflux protein corC	9.B.37
GD1900	surface antigen	1.B.33
GD2004	modE Molybdenum transport protein modE	3.A.1.8
GD2006	TonB-dependent receptor	1.B.14.1
GD2030	Mg chelatase-related protein	3.A.1
GD2065	Probable allantoin permease	2.A.39.3

GD2073	putative ABC transporter ATP-binding protein	3.A.1.120
GD2074	Outer membrane protein	2.A.6.2
GD2075	mdtO Multidrug resistance protein mdtO	8.A.1
GD2076	mdtN Multidrug resistance protein mdtN	8.A.1
GD2092	cyoD Cytochrome o ubiquinol oxidase protein cyoD	3.D.4
GD2093	cyoC Cytochrome o ubiquinol oxidase subunit 3	3.D.4
GD2094	cyoB Ubiquinol oxidase subunit 1	3.D.4
GD2095	cyoA Ubiquinol oxidase subunit 2 precursor	3.D.3
GD2097	putative membrane protein	9.B
GD2116	petC Cytochrome c	3.D.3
GD2117	petB Cytochrome b	3.D.3
GD2121	btuC Vitamin B12 import system permease protein btuC	3.A.1.14
GD2122	putative periplasmic binding protein	3.A.1
GD2124	ccp Cytochrome c551 peroxidase precursor	5.B.3.1
GD2127	putative transporter protein	9.B
GD2128	secB Protein-export protein secB	9.B
GD2135	Inner membrane transport protein	2.A.41
GD2161	ctaB Protoheme IX farnesyltransferase	3.D.4
GD2162	ctaD Cytochrome c oxidase subunit 1	3.D.4
GD2179	Probable ABC transporter ATP-binding protein	3.A.1.4
GD2196	oxaA Inner membrane protein oxaA	2.A.9
GD2206	yaeT Outer membrane protein assembly factor yaeT precursor	1.B.33.1
GD2225	putative membrane protein	2.A.1
GD2230	galP Galactose-proton symporter	2.A.1.1
GD2248	Outer membrane lipoprotein	1.B.6.1
GD2261	norM Multidrug resistance protein norM (Multidrug-efflux transporter)	2.A.66.1
GD2284	tatC Sec-independent protein translocase protein	2.A.64
GD2294	gabP GABA permease	2.A.3.1
GD2302	Permease protein	2.A.1.40
GD2356	ABC transporter, ATP-binding protein	3.A.1
GD2360	Major facilitator superfamily	2.A.1
GD2374	ftsY Cell division protein ftsY	3.A.5.1
GD2381	ABC-type Fe ³⁺ transport system ATP-binding component	3.A.1
GD2382	ABC-type Fe ³⁺ transport system permease component	3.A.1
GD2383	ABC-type Fe ³⁺ transport system periplasmic component	3.A.1
GD2391	uvrA UvrABC system protein A	3.A.1
GD2411	amiA Probable N-acetylmuramoyl-L-alanine amidase amiA	9.B
GD2415	putative membrane protein	9.B
GD2423	nrgA Ammonia channel precursor (Ammonia transporter)	2.A.49
GD2432	zitB Zinc transporter zitB	2.A.4.1.4
GD2433	fieF Cation-efflux pump fieF	2.A.4.1
GD2434	clcA H(+)/Cl(-) exchange transporter clcA	1.A.11
GD2457	kup potassium transport system protein	2.A.72.1
GD2463	Major facilitator transporter protein	2.A.1.2
GD2482	ctrD Capsule polysaccharide export ATP-binding protein ctrD	3.A.1
GD2483	ctrC Capsule polysaccharide export inner-membrane protein	8.A.4
GD2484	ctrB Capsule polysaccharide export inner-membrane protein	8.A.4
GD2486	ctrA Putative polysaccharide export protein	1.B.18
GD2487	Inward rectifier potassium channel	1.A.2.1
GD2503	Lysine exporter protein	2.A.75
GD2508	putative inner membrane protein	2.A.1
GD2513	secretion protein HlyD	2.A.6
GD2514	Cobalt-zinc-cadmium resistance protein	2.A.6.1.2

GD2515	efflux pump outer membrane protein	2.A.6.2
GD2517	rbsC Ribose transport system permease	3.A.1.2
GD2518	rbsB D-ribose-binding protein precursor	3.A.1.2
GD2528	fbpC Ferric cations import ATP-binding protein fbpC	3.A.1
GD2536	nuoN NADH-quinone oxidoreductase chain N	3.D.1
GD2537	nuoM NADH-quinone oxidoreductase chain M	3.D.1
GD2538	nuoL NADH-quinone oxidoreductase chain L	3.D.1
GD2539	nuoK NADH-quinone oxidoreductase chain K	3.D.1
GD2540	nuoJ NADH-quinone oxidoreductase chain J	3.D.1
GD2541	nuoI NADH-quinone oxidoreductase subunit I	3.D.1
GD2542	nuoH NADH-quinone oxidoreductase subunit H	3.D.1
GD2543	nuoG NADH-quinone oxidoreductase chain 3	3.D.1
GD2544	nuoF NADH-quinone oxidoreductase chain F	3.D.1
GD2545	nuoE NADH-quinone oxidoreductase chain E	3.D.1
GD2546	nuoCD ADH-quinone oxidoreductase chain C/D	3.D.1
GD2547	nuoB NADH-quinone oxidoreductase chain B	3.D.1
GD2548	nuoA NADH-quinone oxidoreductase chain A	3.D.1
GD2565	putative membrane protein	9.B
GD2571	Succinoglycan biosynthesis transport protein	8.A.3
GD2577	Putative O-antigen transporter	2.A.66
GD2590	putative small multidrug resistance protein	2.A.7
GD2595	Sodium/substrate symport protein	2.A.21
GD2596	putative membrane protein	9.B
GD2601	putative membrane protein	9.B
GD2606	Putative proline/betaine transporter	2.A.1.6
GD2609	major facilitator family transporter	2.A.1
GD2613	xylH Xylose transport system permease protein xylH	3.A.1.2
GD2614	xylG D-xylose transport ATP-binding protein xylG	3.A.1.2
GD2615	xylF D-xylose-binding periplasmic protein precursor	3.A.1.2
GD2616	gumJ Lipopolysaccharide biosynthesis protein	2.A.66
GD2648	major facilitator superfamily (MFS) transporter	2.A.1
GD2653	amino acid permease	2.A.3.1.9
GD2662	kefC Glutathione-regulated potassium-efflux system protein	2.A.37.1
GD2675	potD Spermidine/putrescine-binding periplasmic protein	3.A.1.11
GD2678	putative membrane protein	9.B
GD2679	ctaE Cytochrome c oxidase subunit 3	3.D.4
GD2680	ctaD Cytochrome c oxidase subunit 1	3.D.4
GD2681	ctaC Cytochrome c oxidase subunit 2 precursor	3.D.4
GD2706	4Fe-4S ferredoxin, iron-sulfur binding	5.A.3
GD2710	putative transporter protein	2.A.76
GD2712	rbsB D-ribose-binding protein precursor	3.A.1.2
GD2713	rbsC Ribose transport system permease protein rbsC	3.A.1.2
GD2714	rbsA Ribose transport ATP-binding protein rbsA	3.A.1.2
GD2718	rbsC Ribose transport system permease protein	3.A.1.2
GD2719	ABC transporter, ATP-binding protein	3.A.1.2
GD2721	putative transporter protein	3.A.1.1
GD2722	putative transporter protein	3.A.1.1
GD2723	putative transporter protein	3.A.1.1
GD2724	putative sugar uptake ABC transporter ATP-binding protein	3.A.1.1
GD2727	oprB Porin B precursor	1.B.19.1
GD2728	rbsC Ribose transport system permease protein	3.A.1.2
GD2730	nrtA Nitrate transport protein nrtA	3.A.1.16
GD2731	nrtB Nitrate transport permease protein nrtB	3.A.1.16

GD2732	nrtC Nitrate transport ATP-binding protein nrtC	3.A.1.16
GD2741	metQ D-methionine-binding lipoprotein metQ	3.A.1.24
GD2742	metN D-methionine transport ATP-binding protein metN	3.A.1.24
GD2743	metI D-methionine transport system permease protein	3.A.1.24
GD2744	fucl L-fucose isomerase	3.A
GD2745	rbsB D-ribose-binding periplasmic protein precursor	3.A.1.2.1
GD2746	rbsC Ribose transport system permease protein	3.A.1.2.1
GD2762	putative membrane protein	9.B
GD2780	Membrane lipoprotein	3.A.1.2
GD2781	ABC transporter, ATP-binding protein	3.A.1.2
GD2782	putative amino acid transporter	3.A.1.2
GD2783	putative amino acid transporter	3.A.1.2
GD2789	TonB-dependent receptor	1.B.14.1
GD2792	TonB-dependent receptor	1.B.14.1
GD2797	ABC transporter ATP-binding protein	3.A.1.3
GD2798	putative membrane protein	3.A.1.3
GD2800	Inner membrane amino-acid ABC transporter permease	3.A.1.3
GD2801	artJ Arginine-binding periplasmic protein 2 precursor	3.A.1.3
GD2809	TonB-dependent receptor	1.B.14
GD2828	trbI conjugation TrbI-like protein	3.A.7
GD2830	trbF Conjugal transfer protein trbF	3.A.7
GD2832	trbL conjugal transfer protein; TrbL	3.A.7
GD2834	trbE putative conjugal transfer protein;	3.A.7
GD2836	trbC Conjugal transfer protein trbC	3.A.7
GD2837	Type II/IV secretion system protein	3.A.7
GD2841	traG Conjugal transfer protein traG	3.A.7.1
GD2874	Copper-transporting P-type ATPase	3.A.3.5
GD2875	clcA H(+)/Cl(-) exchange transporter	1.A.11
GD2927	nonribosomal peptide synthetases (NPRS)	1.D.1
GD2929	major facilitator superfamily (MFS) transporter	2.A.1
GD2944	putative multidrug resistance protein	2.A.1.3
GD2945	putative multidrug resistance protein	2.A.1.3
GD2946	multidrug efflux MFS outer membrane protein	2.A.6.2
GD2961	traF putative conjugal transfer protein traF	3.A.7
GD2995	putative multidrug resistance protein	2.A.1
GD3001	trbI putative conjugal transfer protein trbI	3.A.7
GD3003	CagX putative conjugal transfer protein	3.A.7
GD3004	trbF putative conjugal transfer protein	3.A.7
GD3005	trbL TrbL/VirB6 putative conjugal transfer protein	3.A.7
GD3007	trbE putative conjugal transfer protein trbE	3.A.7
GD3008	trbD putative conjugal transfer protein trbD	3.A.7
GD3009	trbC putative conjugal transfer protein	3.A.7
GD3010	trbB conjugal transfer protein trbB	3.A.7
GD3013	traG Conjugal transfer protein traG	3.A.7
GD3028	putative permease	2.A
GD3029	putative permease protein	2.A
GD3053	Putative sulfate transporter	2.A.53
GD3055	ABC transporter substrate-binding protein	3.A.1
GD3057	Putative ABC transport system permease protein	3.A.1
GD3112	putative amino acid permease protein	2.A.3.3
GD3121	loID Lipoprotein-releasing system ATP-binding protein loID	3.A.1
GD3122	loIC Lipoprotein-releasing system transmembrane protein	3.A.1
GD3127	nouN NADH-quinone oxidoreductase	3.D.1

GD3128	nouL NADH-quinone oxidoreductase	3.D.1
GD3129	nouK NADH-quinone oxidoreductase	3.D.1
GD3130	nouJ NADH-quinone oxidoreductase	3.D.1
GD3131	nouI NADH-quinone oxidoreductase chain I	3.D.1
GD3132	nouH NADH-quinone oxidoreductase	3.D.1
GD3133	nouG NADH-quinone oxidoreductase	3.D.1
GD3134	nouF NADH-quinone oxidoreductase	3.D.1
GD3135	nouE NADH-quinone oxidoreductase	3.D.1
GD3136	nouD NADH-quinone oxidoreductase	3.D.1
GD3137	nouC NADH-quinone oxidoreductase	3.D.1
GD3138	nuoB NADH-quinone oxidoreductase chain B	3.D.1
GD3139	nouA NADH-quinone oxidoreductase	3.D.1
GD3149	Predicted integral membrane protein	9.B
GD3159	sufD FeS assembly protein SufD	3.A.1
GD3160	sufC FeS assembly protein SufC	3.A.1
GD3161	sufB FeS assembly protein SufB	3.A.1
GD3163	Predicted membrane protein	2.A
GD3170	putative ABC transporter	3.A.1
GD3174	putative membrane protein	8.A.1
GD3181	putative lipoprotein (SmpA/OmlA family)	8.A
GD3195	ffh Signal recognition particle protein	3.A.5
GD3197	phoU Phosphate transport system protein	3.A.1.7
GD3198	pstB Phosphate import ATP-binding protein pstB	3.A.1.7
GD3199	pstA Phosphate transport system permease protein	3.A.1.7
GD3201	pstC Phosphate transport system permease protein	3.A.1.7
GD3205	putative ABC transporter	3.A.1
GD3210	putative oligopeptide transporter	2.A.67
GD3226	putative sugar transporter	2.A.1.1
GD3228	putative thiamine transport ATP-binding protein	3.A.1
GD3235	putative cobalamin binding protein	9.B
GD3236	putative heavy metal transporter	3.A.3.5
GD3237	putative heavy metal transporter	3.A.3
GD3257	putative 4-amino-4-deoxy-L-arabinose transferase	3.A.1
GD3284	Ubiquinol-cytochrome c reductase iron-sulfur subunit	3.D.3
GD3289	Peptidoglycan synthetase ftsI	2.A.7
GD3290	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	2.A.7
GD3291	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	2.A.7
GD3292	Phospho-N-acetylmuramoyl-pentapeptide-transferase	4.A
GD3293	UDP-N-acetylmuramoylalanine--D-glutamate ligase	2.A.7
GD3295	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	2.A.7
GD3296	UDP-N-acetylmuramate--L-alanine ligase	2.A.7
GD3297	UDP-N-acetylenolpyruvoylglucosamine reductase	2.A.7
GD3305	putative membrane protein	9.B
GD3313	putative ABC transporter	3.A.1
GD3315	putative amino acid transporter	3.A.1
GD3316	putative amino acid transporter	3.A.1
GD3321	putative membrane protein	2.A.1
GD3327	putative transporter protein	2.A.1
GD3329	putative ABC transporter	3.A.1
GD3330	putative ABC transporter	3.A.1
GD3355	Outer membrane lipoprotein carrier protein LolA	1.B
GD3361	putative transporter protein	2.A.1.3

GD3363	Galactose-proton symporter	2.A.1.1
GD3365	Gluconate permease	2.A.8
GD3389	putative transporter protein	2.A.1.3
GD3393	two-component sensor histidine kinase/response regulator hybrid protein	2.A.21
GD3394	putative histidine kinase	2.A.21
GD3401	putative membrane protein	9.B
GD3402	putative ABC transporter	3.A.1
GD3417	NADH-quinone oxidoreductase chain B	3.D.1
GD3420	putative membrane protein	2.A
GD3421	putative membrane protein	9.B
GD3422	putative membrane protein	9.B
GD3423	probable Na ⁺ -translocating NADH-quinone dehydrogenase	3.D.5
GD3424	putative oxidoreductase	8.A.5
GD3431	putative permease	2.A
GD3432	putative permease protein	2.A
GD3433	Organic solvent tolerance protein precursor	1.B.42.1.1
GD3459	Multidrug resistance protein B	2.A.1.3
GD3460	Multidrug resistance protein A	2.A.1.3
GD3461	putative ABC transporter protein, ATP-binding cassette	3.A.1.6.1
GD3475	Multidrug resistance protein mdtA	2.A.6.2
GD3476	Multidrug resistance protein mdtB	2.A.6.2
GD3477	Multidrug resistance protein mdtC	2.A.6.2
GD3478	Putative outer membrane drug efflux protein	2.A.6.2
GD3484	hypothetical membrane protein	9.B
GD3495	Preprotein translocase secY subunit	3.A.5.1.1
GD3512	putative secretion protein	8.A.1
GD3513	putative ABC transporter protein	3.A.1.109
GD3522	putative autotransporter protein	1.B.12
GD3525	conserved hypothetical protein	89.B.20
GD3527	putative outer membrane efflux protein	1.B.17
GD3528	putative carbohydrate-selective porin	1.B.19
GD3528	putative transporter protein	2.A.6.2
GD3529	acrA putative acriflavin resistance protein	2.A.6.2
GD3530	putative two-component sensor histidine kinase	2.A.21
GD3530	putative outer membrane protien	2.A.6
GD3545	putative peptide ABC transport protein	3.A.1.5
GD3546	putative peptide ABC transport protein	3.A.1.5
GD3547	putative transport/binding protein	3.A.1.5
GD3548	cycZ heme exporter protein C	3.A.1.107
GD3548	putative transport/binding protein	3.A.1.5
GD3549	putative peptide bind protein	3.A.1.5
GD3552	cycK Cytochrome c-type biogenesis protein	9.B.14
GD3559	putative membrane protein	9.B
GD3568	putative membrane protein	9.B
GD3569	putative transporter protein	2.A.1
GD3570	putative TonB-dependent receptor	1.B.14
GD3589	putative membrane protein	9.B
GD3602	ABC transporter ATP-binding protein	3.A.1.120
GD3605	putative membrane protein	9.B
GD3753	putative TonB-dependent Receptor protein	1.B.14.1.2
GD3756	putative sulfate trasporter	2.A.53.4.1
GD3765	Carbohydrate-selective porin	1.B.19.1.1
GD3772	Arsenical pump membrane protein	3.A.4.1.1

GD3803	putative amino acid permease	2.A.3.1
GD3812	putative Outer membrane protein	2.A.6.2
GD3823	putative TonB-dependent receptor	1.B.14.1
GD3824	putative tripeptide permease	2.A.17
GD3829	putative phosphate transport system permease protein pstC	3.A.1
GD3830	putative phosphate transport system permease protein pstA	3.A.1
GD3831	Phosphate import ATP-binding protein pstB	3.A.1
GD3835	putative phosphate-binding periplasmic protein precursor	3.A.1.7
GD3836	putative TonB-dependent Receptor protein	1.B.14.1.4
GD3838	putative TonB-dependent Receptor protein	1.B.14.1.3
GD3840	putative phosphate transporter	2.A.20.1
GD3842	putative TonB-dependent Receptor protein	1.B.14.1.3
GD3843	putative Outer membrane protein	2.A.6.2.23
GD3844	putative ABC transporter	3.A.1.3
GD3845	putative transport protein	3.A.1.3
GD3846	putative ABC transport system, ATP-binding protein	3.A.1.2.9
GD3850	putative transport ATP-binding protein	3.A.1.2.1
GD3852	putative transport system permease protein	3.A.1.2.9
GD3853	putative transport system permease protein	3.A.1.2.8
GD3854	conserved hypothetical protein	3.A.1.2.9
GD3886	Secretion protein HlyD	8.A.1.1.1
GD3887	putative Outer membrane protein oprM	2.A.6.2.23
GD3888	putative chloride channel protein	1.A.11.6.1
GD3967	putative outer membrane efflux protein	2.A.6.2.21
GD3969	putative secretion protein HlyD	2.A.6.2.20
GD3972	putative C4-dicarboxylate transport protein	2.A.23.1.2
GD3975	Multidrug resistance protein mdtB	2.A.6.2.12
GD3976	putative secretion protein HlyD	2.A.6.2.20
	putative outer membrane factor	1.B.17
	putative Na(+)/H(+) antiporter	2.A.36.3.2
	putative efflux transporter, RND family. Pseudogene	2.A.6
	putative heavy metal efflux protein. Pseudogene.	2.A.6.1
	putative ABC transporter component	3.A.1
	putative conjugal transfer protein trbG/irB9	3.A.7.1.1
	putative conjugal transfer protein trbI/irB10	3.A.7.1.1
	putative conjugal transfer protein	3.A.7.1.1
	putative conjugal transfer protein	3.A.7.1.1
	putative conjugal transfer protein	3.A.7.1.1
	putative conjugal transfer protein	3.A.7.1.1

2.3.1 Visão geral dos transportadores.

Após a identificação das proteínas de transporte presentes no genoma de *G. diazotrophicus* sua quantidade total e porcentagem relativa ao número total de fases abertas de leitura (ORFs) foram comparadas a algumas bactérias que tiveram seus genomas divulgados (Tabela 2.2).

Tabela 2.2: Comparação entre o sistema de transportes no genoma da *G. diazotrophicus* com outros genomas publicados

Organismos	Proteobacteria	Tamanho do Genoma (Mb)	Total ORFs	Total de Transportadores	No. de Transportadores por Mb genoma	Total Transportadores como Percentagem de ORFs Totais
<i>A. tumefaciens</i>	Alpha	5,3	5419	357	69,25	6,6
<i>B. japonicum</i>	Alpha	9,11	8317	550	60,37	6,6
<i>M. loti</i>	Alpha	7,59	6752	883	57,05	13,1
<i>S. meliloti</i>	Alpha	6,7	6204	826	57,76	13,3
<i>G. oxydans</i>	Alpha	2,92	2664	146	50,00	5,5
<i>G. diazotrophicus</i>	Alpha	4,3	3854	563	130,93	14,6
<i>R. prowazekii</i>	Alpha	1,1	835	48	43,63	5,7
<i>Azoarcus</i> sp.	Beta	4,74	3992	200	42,1	5
<i>C. violaceum</i>	Beta	4,7	4431	489	104,0	11
<i>X. fastidiosa</i>	Gamma	2,83	2904	78	29,1	2,7
<i>E. coli</i> K-12	Gamma	4,6	4288	354	76,96	8,3

Desse modo, identificamos que os sistemas de transportes representam 14,6% dos genes no genoma *G. diazotrophicus*. Vale ressaltar que as espécies do grupo alfa-Proteobacteria exibem grandes diferenças correspondentes aos índices de transportadores que podem ser decorrentes da variedade de estilos de vida adotadas por essas bactérias. Por exemplos: organismos relacionados a rizosferas, tais como *Sinorhizobium meliloti* e *Mesorhizobium loti* apresentam respectivamente 13,3% e 13,1% de suas proteínas envolvidas com sistemas de transporte, já alguns patógenos intracelulares obrigatórios ou simbiontes, apresentam um número muito menor de transportadores, como por exemplo a *Rickettsia prowazekii* que apresenta apenas 5,7% de suas proteínas como transportadores. Já micróbios associados à planta e ao solo codificam um robusto sistema de transportadores, refletindo a versatilidade destes organismos à exposição frente uma larga escala de substratos diferentes em seus ambientes naturais [Ren e Paulsen, 2005].

Um olhar mais apurado para o sistema de transporte é vital à compreensão da potencialidade metabólica de organismos completamente seqüenciados. Entretanto, é frequentemente problemático efetuar tais anotações, por causa da ocorrência de famílias grandes e complexas de transportadores, como as transportadores ABC e superfamílias MFS, e a presença de múltiplos parálogos em muitos organismos [Ren e Paulsen, 2005]. Nos utilizamos o banco de dados do TCDB (Transport Classification Database) para efetuar a classificação das ORFs. Na tabela 2.3 são apresentadas as principais classes de

transportadores identificadas no genoma de *G. diazotrophicus*

Tabela 2.3: Principais classes de transportadores anotados no genoma de *G. diazotrophicus*

Categoria de transportador	Número de ORFs anotadas	% em relação a ORFs de transportadores anotadas	% em relação a ORFs totais anotadas
Transportadores primários ativos	257	45,6	6,4
Transportadores secundários	167	29,7	4,2
Poros e canais	68	12,1	1,7
Outros	71	12,6	1,8
TOTAL	563	100	14,1

2.3.2 Transportadores primários

O maior grupo de transportadores anotados (257 ORFs) codifica para transportes efetuados por **transportadores primários** (45,6%), dos quais 130 são transportadores do tipo ABC. Transportadores do tipo ABC constituem o principal sistema de transporte em *G. diazotrophicus* (~23,1% de todas ORFs ligadas ao sistema de transporte). A maioria (~70%) das ORFs de transportadores do tipo ABC em *G. diazotrophicus* são relacionadas à aquisição de nutrientes, tais como: fosfatos, fosfonatos, sulfato/molibdato, aminoácidos (glutamato/aspartato, leucina/isoleucina/valina, arginina/ornitina, taurina, histidina), metais (cobre, ferro, magnésio, manganês/zinco, níquel, potássio), espermidina/putrescina, dipeptídeos/oligopeptídeos, e açúcares (ribose, xilose, arabinose, β -galactose e glicerol-3 fosfato), ao passo que as restantes estão relacionadas com a multi-resistência. Ressalta-se ainda a presença de um transportador de nitrato (GDI1817). A grande maioria destes transportadores do tipo ABC estão organizados em genes agregados (que provavelmente funcionam como operons) composto por três componentes: um componente ATPase e duas proteínas auxiliares.

Entre as proteínas envolvidas no metabolismo energético, uma ATP sintase (*atpD*) foi identificada. Dados de proteômica de *G. diazotrophicus* PAL5 em condições de fixação de nitrogênio mostram a expressão diferencial dessa *atpD* por *G. diazotrophicus* na fase L (logarítmica) de crescimento sem controle de pH. Este fato é interessante e pode fazer parte de uma estratégia envolvendo o sistema Kdp (mencionado no item 2.3.4), para manter a homeostase citoplasmática e o pH interno (pHi) [Lery et al., 2008].

Foram encontrados ainda 41 genes de secreção tipo IV que estão relacionados com a exportação lipooligosacarídeos e sistemas de conjugação (transporte de proteínas, DNA e outras macromoléculas). Isso talvez tenha permitido a aquisição de uma grande quantidade de genes de transferência laterais, como também a presença de 38 genes responsáveis pela produção de glicosiltransferases, que são genes requeridos para biossíntese de polisacarídeos de superfície celular (principalmente succinoglicanos). Acredita-se que estes apresentem importante papel de reconhecimento celular, durante a interação bactéria-planta os quais correspondem a requisitos necessários ao reconhecimento da bactéria pela planta, e conseqüente estabelecimento da simbiose [Krause et al., 2006].

2.3.3 Transportadores secundários

O segundo maior grupo de transportadores anotados (167 ORFs) codifica para transportes efetuados por **transportadores secundários** (29,7%). Neste grupo encontram-se as importantes famílias MSF (“Major Facilitator Superfamily”), RND (Resistance Nodulation Division) e TonB. Foram identificados no genoma de *G. diazotrophicus* 48 transportadores da superfamília MFS. A MFS é uma superfamília muito antiga, ampla e diversificada que inclui mais de 10.000 membros seqüenciados. Seu transporte pode ser do tipo uniporte, simporte soluto:cátion (H^+ ou Na^+) ou antiporte soluto: H^+ . Transportadores MFS apresentam especificidade para açúcares, polióis, drogas, neurotransmissores, metabólitos do ciclo de Krebs, intermediários fosforilados da via glicolítica, aminoácidos, peptídeos, osmolites, nucleosídeos, ânions orgânicos, ânions inorgânicos, entre outros.

Além disso, foram identificados 40 genes da família RND. Esse sistema utiliza a força

motriz de prótons como fonte de energia, eles existem em todos os reinos de organismos vivos, e parecem estar envolvidos na resistência a drogas, especialmente em bactérias Gram negativas [Paulsen et al., 1996]. Os transportadores RND funcionam como um complexo de proteínas que vão desde a membrana citoplasmática até a membrana externa.

Ressalta-se ainda a ocorrência de 23 genes transportadores do tipo TonB e 6 receptores dependentes de TonB no genoma de *G. diazotrophicus*. Os receptores dependentes de TonB são uma família de proteínas da membrana externa de bactérias gram-negativas. O complexo TonB recebe os sinais de fora da célula bacteriana e transmite-os para o meio intracelular, levando a ativação transcricional de genes alvo. Em *Escherichia coli* proteínas TonB interagem com proteínas receptoras da membrana exterior que realizam a ligação por alta afinidade e a captação (dependente de energia) de substratos específicos, como cobalamina, colicinas e principalmente Ferro [Chimento et al., 2003]. Na ausência de TonB, esses receptores se ligam a seus substratos, mas não realizam transporte ativo [Koebnik, 2005]. O Ferro em geral é muito pouco permeável através de canais porina, além de ser encontrado em concentrações extracelulares muito baixas. No entanto é um elemento essencial para muitas proteínas, entre eles o complexo nitrogenase, citocromos e oxidoreductases, relacionadas ou não com cadeia transportadores de elétrons. A presença de uma grande quantidade destes destes transportadores no genoma da *G. diazotrophicus* pode ser essencial para o caráter diazotrófico desta bactéria [Rosconi et al., 2006].

Além disso, genes outros experimentos de proteômica mostram que uma proteína hipotética que é superexpressa quando crescida em contato com a cana-de-açúcar tem um peptídeo sinal e está dentro de um operon importante para o transporte de biopolímero em *Gluconobacter oxydans* 621H. Este operon inclui EXBD e ExBB, proteínas do sistema TonB [dos Santos et al., 2010]. Em estudos anteriores cujo objetivo era estudar os mecanismos moleculares que controlam a adaptação das bactérias fitopatogênicas às suas plantas hospedeiras, foi realizada a exploração de sequências genômicas completas de 226 bactérias Gram negativas. Esses estudos mostraram que a representação de receptores dependentes de TonB é restrita a uma pequena proporção dessas bactérias. A

maioria das bactérias que exibem essa particularidade têm estilos diferentes e pertencem a diferentes linhagens taxonômicas, mas todas elas compartilham a capacidade de explorar carboidratos complexos. Assim, foi postulado que alguns receptores dependentes de TonB podem estar envolvidos no transporte de moléculas derivadas de plantas, principalmente carboidratos como sacarose [Blanvillain et al., 2007].

2.3.4 Poros e canais

Foram identificadas ainda 68 ORFs que codificam para **transportadores do tipo “canal ou poro”**. Entre elas podemos ressaltar a presença de uma K^+ histidina quinase (*KdpD*). Em *E. coli* e outras bactérias, KdpD/KdpE são, respectivamente, o sensor quinase e o sistema regulador da expressão do operon *kdpFABC*, que codifica um transportador de alta afinidade a K^+ . Experimentos de proteômica mostram a expressão diferencial (superexpressão) desta uma proteína em condições de fixação de nitrogênio durante a fase estacionária (fase S) de crescimento. Em *G. diazotrophicus*, tal sistema poderia ser uma estratégia para manter a homeostase citoplasmática e controle do pHi (pela extrusão de H^+ durante o influxo de K^+) no meio ácido em que a bactéria sobrevive [Lery et al., 2008].

Foram identificadas ainda duas proteínas similares a transportadores de NH_3 ou NH_4^+ do tipo Amt (GD 0731 e GD2352). As proteínas da família Amt parecem possuir pelo menos duas funções celulares: transportando NH_3 ou NH_4^+ e regulando o metabolismo de nitrogênio se ligando diretamente a proteínas do tipo PII e seu homólogo GlnK. Normalmente amtB e glnK formam um operon, e GlnK regula a atividade de AmtB. Em *Azospirillum brasilense*, AmtB forma um complexo com GlnZ e a atividade reguladora da nitrogenase DraG [Huergo et al., 2007]. Em *G. diazotrophicus* também é possível observar a formação do operon composto por amtB (GD2352) e glnK (GD2352), o que pode indicar que essa proteína poderia ter um papel importante no metabolismo de nitrogênio desta bactéria.

2.4 Conclusões

Os dados apresentados aqui mostram os resultados da anotação dos transportadores de *G. diazotrophicus*. Nesta etapa do trabalho foram analisadas 1160 sequências codificantes para proteínas de membrana e 563 foram confirmadas como proteínas transportadoras. Cerca de 45% destes transportadores anotados pertencem à classe de “transportadores ativos primários”, 30% à classe de “transportadores ativos secundários”, e 12% relacionadas à formação de poros e canais na membrana celular. 13% deles não pertencem a nenhum destes grupos ou não puderam ter uma caracterização mais precisa.

A análise da anotação dos transportadores genoma de *G. diazotrophicus* já permite uma primeira caracterização do perfil metabólico desta bactéria. Deste tipo de análise podemos ressaltar:

- A presença de 23 genes transportadores do tipo TonB que estão ligados à captação de ferro extracelular e 6 receptores dependentes de TonB, que podem estar envolvidos na utilização de carboidratos vegetais pela bactéria parecem mostrar uma grande adaptação ao seu perfil de bactéria endofítica fixadora de nitrogênio.
- A presença de 40 genes da família RND e uma grande quantidade de genes para exportar metais pesados e outros compostos tóxicos, corroborada pelo percentual de transportadores ativos no genoma, concedem à bactéria uma notável resistência a ambientes inóspitos.
- A presença de transportadores de um sistema de secreção do tipo IV confere à bactéria a capacidade de transportar macromoléculas através de seu envelope até a planta. Acredita-se que isso tenha um papel relevante no reconhecimento celular, no curso da interação planta-bactéria.
- A presença de um canal osmosensitivo de K^+ histidina quinase (KdpD) e de uma ATP sintase (*atpD*) que podem estar envolvidas na manutenção da homeostase citoplasmática e controle do pH interno.

- A presença de um operon formado por *amtB* e *glnK* o que indicar que essa proteína poderia ter um papel crucial no metabolismo de nitrogênio desta bactéria.

Capítulo 3

As espécies reativas de oxigênio durante a FBN

3.1 Introdução

3.1.1 O paradoxo da FBN em organismos aeróbicos

A FBN é considerada um processo extremamente dispendioso para a célula pois além do poder redutor necessário para a realização da reação, esta requer o gasto de 16 moléculas de *ATP* para cada molécula de N_2 reduzida (equação 1.1). Essa quantidade de *ATP* pode ser suprida satisfatoriamente pela respiração aeróbica, mas há um problema nisso: as duas proteínas que compõem o complexo da nitrogenase, a dinitrogenase reductase (Fe-proteína) e a dinitrogenase (MoFe-proteína) (Figura 1.1), são inativadas irreversivelmente por oxigênio [Ureta e Nordlund, 2002]. Essa inibição pode ocorrer de três formas [Goldberg et al., 1987]:

- a nível genético atuando sobre a síntese da nitrogenase;
- causando danos irreversíveis no centro férrico da Fe-proteína;
- causando a inativação temporária (switch-off) da nitrogenase. Neste caso, a atividade pode retornar após a pressão de oxigênio diminuir.

Conseqüentemente, a realização do processo de FBN por organismos aeróbicos sempre foi encarada como um paradoxo pela comunidade científica e é tema de diversos estudos, a maioria deles focados em entender a regulação da expressão gênica por oxigênio e os mecanismos fisiológicos que protegem a nitrogenase do O_2 .

3.1.2 A tolerância de *G. diazotrophicus* a O_2

Alguns autores ainda dizem que a FBN em *G. diazotrophicus* não é tolerante à presença de O_2 . No entanto, já foi observado em *G. diazotrophicus* que quanto maior a taxa de aeração do meio de cultura, maior a atividade específica da nitrogenase e maior a taxa de crescimento obtida [Flores-Encarnacion et al., 1999]. Pan e Vessey também observaram que as colônias de *G. diazotrophicus* crescidas em meio sólido apresentaram atividade ótima da nitrogenase quando incubadas em concentrações de 20% a 30% de oxigênio [Pan e Vessey, 2001], e Rodés e colaboradores mostram que a atividade ótima de nitrogenase no meio semi-sólido é observada em células crescidas com concentração de O_2 entre 10% e 15% [Rodés, 2002]. Esses resultados passam a impressão de que o processo de FBN em *G. diazotrophicus* é não só tolerante, mas que precisa de altas concentrações de oxigênio para sua realização. Mas por que então alguns autores dizem que a FBN em *G. diazotrophicus* não é tolerante à presença de O_2 ?

Essa classificação ocorre pois quando os níveis de oxigênio são medidos na fase líquida do meio de cultura em vez da parte aérea, a atividade de nitrogenase ótima é observada quando a pressão de oxigênio é de 0,2 kPa, ou 4% [Reis e Döbereiner, 1998, Boddey et al., 1991]. Isso é uma indicação de que durante o processo de FBN as células aumentam suas taxas respiratórias, o que aumenta o consumo de oxigênio dissolvido no meio. O aumento no consumo de oxigênio leva conseqüentemente a uma diminuição de sua concentração no meio de cultura de maneira que o oxigênio não é mais capaz de afetar a nitrogenase, que continua funcionando normalmente. Esse tipo de fenômeno é conhecido como proteção respiratória, foi primeiramente descrito em *Azotobacter chroococcum* [Dalton e Postgate, 1968] e será mais detalhado no Ítem 3.1.3.

Neste momento cabe ressaltar que concentração de O_2 no ambiente intracelular onde a nitrogenase atua é resultado da interação de quatro fatores principais [Loiret et al., 2004]:

- concentração de O_2 no meio ambiente;
- taxa de difusão O_2 do ambiente externo ao local da enzima;
- taxa de utilização de O_2 nas proximidades da nitrogenase;
- a presença de moléculas carreadoras de O_2 (leghemoglobina em leguminosas).

3.1.3 Proteção respiratória

Organismos que fixam o nitrogênio geralmente adotam estratégias diferentes para proteger a nitrogenase do contato com O_2 . A mais estudada delas é a estratégia adotada por alguns gêneros (*Rhizobium*, *Bradyrhizobium*, *Sinorhizobium* e *Azorhizobium*) da família *Rhizobiaceae*: ao infectar a planta, eles disparam uma cascata que leva a formação nódulos nas raízes. As células da bactéria ficam protegidas dentro deste nódulo que tem um ambiente anaeróbico que favorece o processo de FBN, enquanto a planta fornece a energia necessária para as bactérias fixarem nitrogênio [Waters e Emerich, 2000]. Diferentemente de organismos da família *Rhizobiaceae*, a *G. diazotrophicus* coloniza os espaços intercelulares da planta (apoplasma) e não forma nódulos, por isso é sempre exposto a níveis mais elevados de O_2 [Dong et al., 1994]. Isso significa que as estratégias de defesa de *G. diazotrophicus* devem ser diferentes das adotadas por organismos que não são expostos a concentrações tão altas de oxigênio.

Vários autores concordam que a proteção respiratória é um mecanismo que desempenha um papel importante na defesa de *G. diazotrophicus* contra o oxigênio. A “defesa respiratória” é um mecanismo, descrito pela primeira vez no gênero *Azotobacter*, segundo o qual as taxas respiratórias extremamente elevadas de células observadas durante a FBN tem a função de manter a concentração de oxigênio intracelular baixa o suficiente para não afetar os componentes da nitrogenase, permitindo a coexistência na

célula da respiração aeróbia e da atividade da nitrogenase [Dalton e Postagate, 1968]. De fato, Flores-Encarnacion e colaboradores mostram que o microorganismo aumenta sua taxa de respiração, quando crescido em condições limitantes de nitrogênio e cultivado na presença de O_2 , e que isso faz com que a concentração de O_2 no ambiente circundante diminua [Flores-Encarnacion et al., 1999, Pan e Vessey, 2001, Ureta e Nordlund, 2002]. Eles também mostram que a *G. diazotrophicus* muda a composição da sua cadeia de transporte de elétrons durante o processo de FBN. Neste contexto, culturas bem aeradas expressam os citocromos a1 e bb como a principal oxidase terminal. Durante a repressão da atividade diazotrófica, a expressão do citocromo a1 é diminuída drasticamente concomitante ao aparecimento do citocromo bd como principal oxidase terminal [Flores-Encarnacion et al., 1999]. Atividades das oxidases também foram muito superiores em preparações de membrana obtidas a partir de culturas em condições de FBN do que aquelas crescidas em condições de não-fixação de nitrogênio. A taxa respiratória de *G. diazotrophicus* foi determinada entre as mais altas já descritas para bactérias aeróbicas [Flores-Encarnacion et al., 1999, Gonzalez et al., 2006].

Além da proteção respiratória, outros organismos fixadores de nitrogênio, tais como *Rhodobacter encapsulados*, *Azospirillum brasilense*, *Azospirillum lipoferum* e *Azotobacter*, possuem também mecanismos de repressão/desrepressão dependente da ligação com uma proteína que oferece uma proteção conformacional à nitrogenase ao oxigênio. Quando exposta a maiores concentrações de O_2 ocorre a ligação dessa proteína reguladora com a nitrogenase, que fica inativa. Quando os níveis de oxigênio voltam a baixar esse complexo é desfeito e a nitrogenase rapidamente retoma sua atividade. Pan e Vessey postularam que deveria haver em *G. diazotrophicus* um mecanismo semelhante, uma vez que, sob certas condições, a atividade da enzima pode ser recuperada após ser inativada por oxigênio [Pan e Vessey, 2001]. Em 2002, Ureta e Nordlund confirmaram essa suposição, mostrando que existe um mecanismo de proteção conformacional da nitrogenase, semelhante ao descrito para *Azotobacter spp.*, onde existe a formação de um complexo entre a proteína FeSII (Shethna) e a nitrogenase [Ureta e Nordlund, 2002]. Este mecanismo

permite que as células se adaptem a altas concentrações de O_2 sem que a enzima seja danificada. Sob estas condições, se a fonte de carbono no meio permite uma alta taxa de respiração, a atividade da enzima será recuperada, caso contrário, a nitrogenase é degradada após um curto período de tempo inativa. O mecanismo de regulação da ligação desta proteína Shethna com a nitrogenase ainda não é conhecido, mas os autores que sugerem que é provavelmente via a percepção do estado redox da célula [Ureta e Nordlund, 2002].

Apesar de vários autores concordarem que a estratégia de proteção respiratória desempenha um papel importante na proteção da nitrogenase [Flores-Encarnacion et al., 1999, Pan e Vessey, 2001, Ureta e Nordlund, 2002], outras investigações ainda precisam ser conduzidas para determinar as formas de regulação desta via e suas implicações para a célula [Loiret et al., 2004].

3.1.4 Ligação entre sistema respiratório e ROS

O aumento da taxa de respiração celular consegue evitar a inibição da nitrogenase pois liga o oxigênio a átomos de hidrogênio formando água. Esse átomos de hidrogênio vieram da quebra da glicose (ou outros polissacarídeos) e foram transferidos através de uma série de proteínas de membrana, chamadas de cadeia transportadora de elétrons, até chegar ao oxigênio. Durante esse processo alguns elétrons podem escapar da cadeia transportadora e se combinar diretamente com um átomo de oxigênio gerando uma forma parcialmente reduzida de oxigênio, comumente chamada de espécie reativa de oxigênio ou ROS (do inglês Reactive Oxygen Species) [Kowaltowski et al., 2009]. Todos os organismos expostos ao oxigênio produzem espécies reativas de oxigênio no seu metabolismo normal. Mais do que simples subprodutos do metabolismo oxidativo as espécies reativas de oxigênio também podem agir de forma benéfica ao organismo, quando usadas pelo sistema imunitário para atacar agentes patogênicos ou quando atuam como moléculas mensageiras em vias de sinalização celular (também designada sinalização redox).

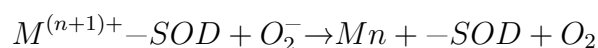
Algumas condições biológicas favorecem a formação de um desequilíbrio entre a produção de ROS e sua desintoxicação através de sistemas biológicos que as removam ou

reparem os danos por elas causados. Esse desequilíbrio é chamado de estresse oxidativo que resulta num ambiente celular mais oxidante [Schafer e Buettner, 2001]. Os efeitos do estresse oxidativo dependem da dimensão de tais variações. Uma célula é normalmente capaz de superar os efeitos do estresse oxidativo se as perturbações no equilíbrio redox forem pequenas, mas perturbações de maior escala podem lesar biomoléculas que, se não puderem ser reparadas pelos sistemas de reparo, levam ao dano celular e tecidual [Santos et al., 2009].

3.1.5 Enzimas detoxificadoras de ROS

Para detoxificar o excesso de ROS geradas pela respiração e manter seu equilíbrio redox, células possuem um grupo de moléculas chamadas de antioxidantes. Os antioxidantes podem ser do tipo “scavenger”, quando agem transformando uma espécie de ROS em outra menos reativa, ou “quencher”, quando consegue neutralizar completamente através da absorção de toda a energia de excitação. Eles também podem ser classificados como enzimáticos e não enzimáticos, conforme a estrutura do agente antioxidante. O sistema enzimático age evitando o acúmulo de ânion radical superóxido (O_2^-) e do peróxido de hidrogênio (H_2O_2). Ele é formado por diversas enzimas, destacando-se a superóxido dismutase (SOD), a catalase (KAT) e a glutathiona peroxidase (GPx).

- A enzima superóxido dismutase (SOD) catalisa a dismutação do superóxido em oxigênio e peróxido de hidrogênio. Ela existe em três formas dependendo do metal associado a ela (Cobre e Zinco no citoplasma de eucariontes, Manganês na matriz mitocondrial e Ferro em bactérias).



- Outro antioxidante enzimático é a catalase, que age sobre o peróxido de hidrogênio transformando-o em água e oxigênio.

- A glutathiona peroxidase (GPx) é uma enzima localizada no citosol e na matriz mitocondrial que reduz o peróxido de hidrogênio e hidroperóxido orgânico utilizando a glutathiona (GSH), um tripeptídeo de ácido α -glutâmico, cisteína e glicina. Esta atua como co-substrato da GPx. Ela é regenerada pela glutathiona redutase (GR) com transferência de hidrogênio do NADPH. Neste processo são transferidos dois hidrogênios dos grupamentos sulfidríla para os peróxidos, transformando-os em álcool e/ou água, resultando em glutathiona oxidada (GSSG). A GPx geralmente ocorre associada ao selênio, mas pode ocorrer independente do mesmo [Halliwell e Gutteridge, 1999].

3.1.6 A influência de ROS na FBN

A deficiência na capacidade de FBN de cepas de *S. meliloti* mutadas em genes envolvidos na homeostase redox indica também o seu papel crucial no metabolismo de FBN: os mutantes afetados na defesa antioxidante (catalase e glutathiona redutase) não foram afetados apenas na sua capacidade de infectar a planta, mas também na eficiência da FBN [Jamet et al., 2003, Harrison et al., 2005]. Da mesma forma, um duplo mutante de *Rhizobium etli* para as enzimas peroxirredoxina (*prxS*) e catalase-peroxidase bifuncional (*katG*) teve sua capacidade de fixação de nitrogênio significativamente reduzida [Dombrecht et al., 2005]. Além disso, cepas de *S. meliloti* mutadas para uma variedade de processos celulares, mas não diretamente envolvidos na defesa antioxidante, também eram afetados em processos de proteção ao estresse oxidativo e estabelecimento da simbiose [Davies e Walker 2007]. De maneira parecida, a mutação de um gene de *S. meliloti* com similaridade a uma tioredoxina envolvido na produção de melanina afetou a resposta a paraquat (estresse oxidativo induzido) e fixação simbiótica de nitrogênio [Castro-Sowinski et al., 2007]. Finalmente, células de *Rhizobium leguminosarum* deficientes na produção de uma tioredoxina-like (*cycY*) foram incapazes de formar nódulos fixadores de nitrogênio em ervilha, já que o mutante estava com defeito na maturação de todos os citocromos tipo c [Vargas et al., 1994].

Uma das características específicas de nódulos é baixa pressão de oxigênio na zona de

fixação para permitir a fixação de nitrogênio eficiente. A regulação da respiração e da eficiência da barreira de difusão do oxigênio desempenham um papel fundamental na proteção do metabolismo do nódulo. Essa regulação parece ser um processo controlado pela planta [Wei e Layzell, 2006]. A diminuição da atividade de FBN induzida por estresse é geralmente associada a uma modificação na pressão de oxigênio no nódulo. Em soja, o estresse salino e hídrico causa uma diminuição na atividade da nitrogenase que pode ser correlacionado com uma diminuição da permeabilidade do nódulo [Del Castillo et al., 1994, Serraj et al., 1994]. Do mesmo modo, os nódulos de soja apresentaram respiração o tamanho do espaço intercelular alterados durante a refrigeração, resultado de alterações na barreira de difusão de oxigênio [van Heerden et al., 2008]. Em contraste, a sensibilidade à salinidade parece estar associada com um aumento na condutância nódulo e um aumento da respiração dos nódulos em *M. truncatula* [Aydi et al., 2004]. Em conjunto, estes resultados mostram que a eficiência do nódulo é fortemente afetada pela regulação da disponibilidade de oxigênio que pode estar ligadas à presença de ROS [Chang et al., 2009].

3.1.7 Modulação de ROS durante a simbiose

ROS são produzidas constantemente pelas plantas como consequência de seu metabolismo aeróbico ou em resposta a estresse biótico ou abiótico. As principais formas de ROS são o oxigênio singlete (O_2), ânion superóxido (O_2^-), peróxido de hidrogênio (H_2O_2), e radical hidroxila ($OH\cdot$). Estas moléculas podem ser altamente tóxicas, uma vez que são capazes de modificar todos os principais constituintes da célula, tais como lipídios, DNA, carboidratos e proteínas [Moller et al., 2007]. Sua toxicidade leva a senescência e morte celular [Overmyer et al., 2003, Rivero et al., 2007]. Por outro lado elas desempenham importantes funções celulares e estão envolvidos na regulação do metabolismo da planta como mensageiros secundários em muitos caminhos associados com o desenvolvimento da planta e respostas ao estresse ambiental [Apel e Hirt, 2004, Carol e Dolan, 2006, Gechev et al., 2006]. Cada vez mais mostra-se que a produção de ROS não é simplesmente uma consequência do metabolismo aeróbico, mas pode ser controlada e modulada, possi-

bilitando que as plantas se adaptem melhor ao ambiente [Pitzschke et al., 2006].

Acumulação de ROS tem sido observada em *Medicago sativa* (alfafa) durante a interação entre rizóbios e leguminosas (Santos et al., 2001). Esse acúmulo de ROS, detectado 12h após tratamento com fatores Nod (NFs), está ligada a uma via de sinalização disparada por NFs. Mutantes de *S. meliloti* que produzem NFs alterados ou um mutante de *Medicago truncatula* (não-nodular), *dmi1-1*, são incapazes de provocar a produção de ROS (Ramu et al., 2002).

Recentemente, mudanças rápidas e transitórias na quantidade de ROS foram observadas nas “root hair cells” minutos após o tratamento com o NFs em *Phaseolus vulgaris* [Cardenas et al., 2008]. A inibição da produção ROS impede “curling” da “root hair cells” e a formação de processos de infecção [Peleg-Grossman et al., 2007]. Assim, a produção de ROS pode não ser uma resposta de defesa contra o micróbios, mas sim um processo que é necessário para o desenvolvimento de uma boa interação. A importância da produção de ROS foi confirmada em uma cepa de *S. meliloti* superexpressando uma catalase. Esta bactéria, atuando como um dissipador de H_2O_2 , provocou um atraso nodulação e a ampliação de processos de infecção [Jamet et al., 2007].

3.1.8 Perspectivas

Os dados analisados nesta revisão indicam claramente que ROS e defesa antioxidante desempenham um papel crucial na simbiose entre leguminosas e rizóbios. O envolvimento de espécies reativas e defesa antioxidante em etapas-chave da formação de nódulos, tais como o desenvolvimento da infecção e formação do meristema do nódulo, mostra que a regulação redox é importante para o desenvolvimento de nódulos. Além disso, várias cepas bacterianas deficientes em defesa antioxidante apresentam uma capacidade de FBN diminuída. Finalmente, modificações na defesa antioxidante da planta são observadas em paralelo com alterações metabólicas do nódulo. Assim, a homeostase redox pode desempenhar um papel crucial no controle do metabolismo de nódulos pela planta. Embora estes estudos já indiquem alguma influência das ROS no processo de FBN, eles foram

realizados com organismos fixadores da família *Rhizobiaceae*. Estes não são bons modelos para a observação da influência direta das ROS na FBN pois só são capazes de fixar nitrogênio quando numa simbiose bem sucedida com a planta. Sabendo-se que a produção de ROS é necessária para o estabelecimento da simbiose pode-se dizer que qualquer alteração no processo de FBN em organismos com o metabolismo redox alterado pode ser uma consequência indireta do não-estabelecimento da simbiose ou da má formação de nódulos.

Além disso, uma caracterização fina da produção de ROS e modulação do estado redox durante a FBN ainda precisa ser alcançada. Chang e colaboradores sugerem que a utilização de sondas específicas, sensíveis ao estado redox ou a acumulação de ROS seriam ferramentas úteis para realizar a análise espaço-temporal da produção de H_2O_2 e alterações redox [Chang et al, 2009]. Neste contexto o objetivo geral desta parte do trabalho foi a caracterização do estado redox das células de *G. diazotrophicus* durante a FBN.

3.2 Objetivos

O objetivo geral desta parte do trabalho foi a caracterização do estado redox das células de *G. diazotrophicus* durante a FBN. Esse objetivo geral pode ser dividido nos seguintes objetivos específicos:

- caracterização do estado redox das células de *G. diazotrophicus* durante a FBN;
- identificação de enzimas responsáveis pela detoxificação de ROS no genoma da bactéria;
- caracterização da atividade da catalase em *G. diazotrophicus*;
- caracterização da expressão de genes de *G. diazotrophicus* que codificam para as enzimas de detoxificação de ROS durante o processo de FBN;
- indicar possíveis alvos para futuros experimentos de mutagenese dirigida visando uma otimização do processo FBN.

3.3 Resultados e Discussão

3.3.1 Visualização

Para testar a hipótese de que o metabolismo de ROS é alterado durante o processo de FBN, a quantidade de ROS foi medida utilizando-se microscopia de fluorescência para analisar células em condições de fixação e não-fixação incubadas com uma sonda sensível a ROS. Para isso células de *G. diazotrophicus* foram crescidas em meio líquido LGI-P suplementado com 1 ou 20 mM de sulfato de amônio. Foi estabelecido que as condições de fixação são aquelas em que a célula é crescida somente com 1 mM de sulfato de amônio no meio de cultura e as condições de não fixação aquelas em que a célula é crescida com 20 mM de sulfato de amônio no meio de cultura. A presença de 20 mM de NH_4 no meio de cultura já é suficiente para inibir a expressão e atividade da nitrogenase (Figura 3.1).

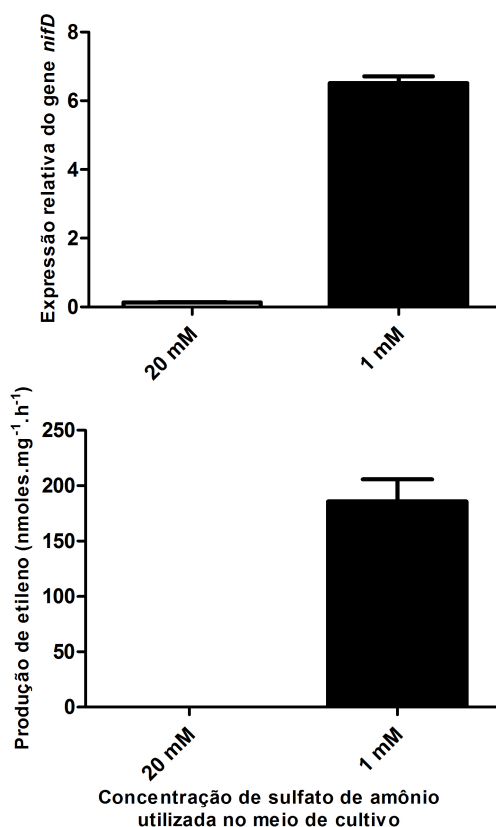


Figura 3.1: Expressão relativa e atividade de redução do acetileno de células de *G. diazotrophicus* crescidas em meio LGI-P suplementado com 1 ou 20mM de sulfato de amônio.

A Figura 3.2 mostra que células crescidas em condições que a estimulem a fixar nitrogênio (FIX) apresentam uma diminuição do sinal de fluorescência (1B e 1D) (indicativo da presença de ROS) quando comparadas a células crescidas em condições de não-fixação (NFIX - 1F e 1H), tanto em 24 quanto em 48h de crescimento.

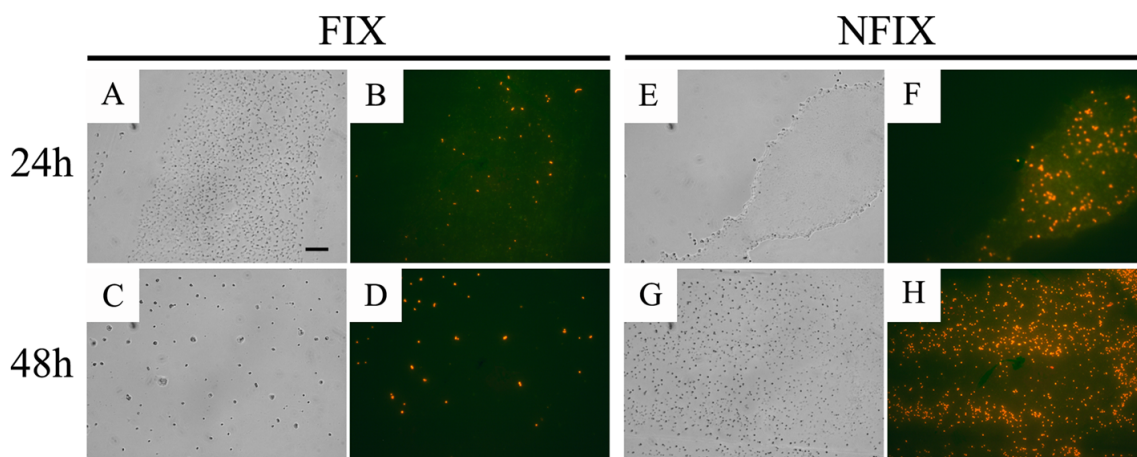


Figura 3.2: Quantidade de ROS é menor durante FBN. Células de *G. diazotrophicus* crescidas em condições de fixação (FIX) e não-fixação de nitrogênio (NFX) foram incubadas com o probe CM-H2DCFDA, sensível a ROS. As Figuras A, C, E e G representam o campo claro das imagens. As Figuras B, D, F e H são microscopias de fluorescência, e os pontos claros indicativos da presença de ROS. Todas as imagens foram obtidas com o mesmo tempo de exposição para permitir a comparação entre a intensidade dos sinais. Barra de escala: 10 μm .

Esses resultados demonstram que a quantidade de ROS intracelular são reduzidos quando *G. diazotrophicus* está fixando nitrogênio, apontando para uma possível correlação inversa entre a presença de ROS e a atividade da nitrogenase. Além disso, é evidente que as células não-fixando habitam um ambiente onde ROS faz parte do meio celular sem exibir qualquer efeito adverso. Isso é consistente com a noção de que ROS agem como moléculas fisiológicas em um conjunto diversificado de situações biológicas [Forman et al., 2010, Terada, 2006], como parece ocorrer em células de *G. diazotrophicus* não fixando nitrogênio. É tentador especular por que os níveis de ROS têm de ser reduzidos durante FBN e nossa hipótese é de que a sensibilidade da nitrogenase à inibição por intermediários de oxigênio é o centro desta questão.

3.3.2 Atividade Catalásica

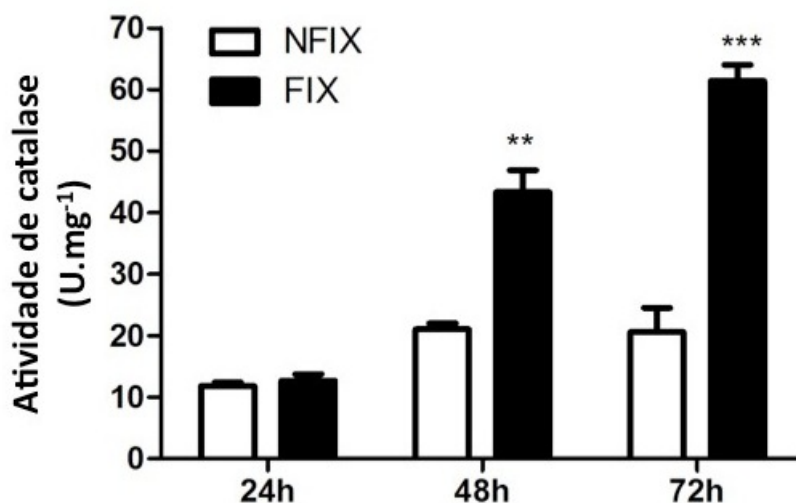


Figura 3.3: Atividade catalásica do extrato bruto de células de *G. diazotrophicus* crescidas em condições de fixação (FIX) e não-fixação de nitrogênio. A atividade catalásica dos extratos brutos celulares foi medida monitorando-se consumo de peróxido de hidrogênio a 240nm por 1 minuto à temperatura ambiente. Média +/- barra de erros de nove experimentos independentes são mostrados.

Para investigar os mecanismos envolvidos da diminuição de ROS intracelular durante o processo de FBN, a atividade catalásica dos extratos brutos celulares foi medida. A catalase é uma enzima antioxidante responsável pela remoção de H_2O_2 . A Figura 3.3 mostra que células de *G. diazotrophicus* crescidas em condições de fixação de nitrogênio apresentam maior atividade catalásica nos tempos de 48 e 72h.

Considerando-se que a atividade de catalase foi aumentada até 6 vezes em células fixando nitrogênio e que, ao mesmo tempo, os níveis de ROS foram grandemente reduzidos (Figura 3.2), resolvemos investigar se outras vias antioxidantes estariam também ativas durante a FBN.

3.3.3 Identificação dos genes responsáveis pela detoxificação de ROS no genoma de *G. diazotrophicus*

Até agora observamos que a quantidade de ROS é menor durante o processo de FBN, quando comparado com culturas não-fixando nitrogênio. A fim de investigar os meca-

nismos envolvidos na defesa desta bactéria contra ROS, foram identificados no genoma recentemente publicado de *G. diazotrophicus* [Bertalan et al, 2009] genes que poderiam estar envolvidos na detoxificação de ROS. A anotação do genoma indica a presença de seis genes que possuem produtos que estão diretamente envolvidos na detoxificação ROS.

Tabela 3.1: Identificação dos genes responsáveis por detoxificação de ROS em *G. diazotrophicus*

Gene	Produto	Locus tag	Seqüencia mais próxima
<i>sodA</i>	Superóxido dismutase	GDI2168	<i>Gluconobacter oxydans</i> (340s, 80%id)
<i>katE</i>	catalase tipo E	GDI0079	<i>Chromobacter salexigens</i> (788s, 74%id)
<i>kat</i>	catalase	GDI0467	<i>Methylobacterium extorquens</i> (332s, 53%id)
<i>katC</i>	catalase tipo C	GDI2359	<i>Azoarcus sp.</i> (1055s, 72%id)
<i>gorA</i>	glutaciona redutase A	GDI2216	<i>Acetobacter pasteurianus</i> (664s, 70%id)
<i>gorB</i>	glutaciona redutase B	GDI2280	<i>Xanthomonas campestris</i> (436s, 63%id)

As primeiro gene analisado foi o da superóxido dismutase (*sodA*, GDI2168), que tem mais homologia com genes de superóxido dismutases de *Alphaproteobacterias* intimamente relacionadas, tais como *Gluconobacter oxydans* 621H (score = 340 bits, identidade = 80%), *Granulibacter bethesdensis* (score = 333 bits, identidade = 79%), e *Acidiphilium cryptum* JF-5 (score = 294 bits, identidade = 74%). A maior parte das ORFs compartilhadas por esses organismos estão relacionados ao metabolismo energético e na transferência de informação e são geralmente parte do núcleo do genoma de *G. diazotrophicus* [Bertalan et al, 2009].

A anotação do genoma de *G. diazotrophicus* também indica a presença de três genes envolvidos na detoxificação ROS no cromossomo principal que codificam para catalases (GDI0079, GDI0467 e GDI2359) [Bertalan et al, 2009]. Dois tipos de catalases estruturalmente independentes são comuns em bactérias: catalase-peroxidases bifuncionais (HPI) e catalases monofuncionais (HPII). Ambas as catalases contêm heme como grupo prostético. Além disso, uma catalase não-heme contendo manganês está presente em algumas

bactérias. A maioria das bactérias parecem manifestar uma ou mais catalases em resposta ao peróxido de estresse, e os diferentes tipos de catalases são regulados de forma independente. Na tabela 3.2 se pode ver os principais tipos de catalase presentes em organismos fixadores de nitrogênio e a classificação das catalases de *G. diazotrophicus*.

A sequência de aminoácidos da ORF GDI0079 (489 resíduos) foi submetida a uma busca por domínios conservados em que identificamos a presença dos domínios "KatE", "heme-binding pocket" e "tetramer interface". Ela apresenta semelhança com catalases monofuncionais de outros organismos estreitamente relacionados, tais como a *katE* de *Chromobacter salexigens* (score = 788 bits, identidade = 74%), *katE* de *Blastopirellula marina* (score = 780 bits, identidade = 75%), e *katE* de *Nitrococcus mobilis* KatE (score = 767 bits, identidade = 73%). Nós chamamos a sequência de *katE*.

A sequência de aminoácidos da ORF GDI0467 (368 residues) foi submetida a uma busca por domínios conservados em que confirmou-se que pertencia à superfamília de catalase. No entanto, a busca por domínios conservados não mostrou a presença de um "heme-binding pocket", um indicativo de que esta catalase não pertence ao tipo HPI ou HPII. Ela tem semelhança com outras catalases, como as de *Methylobacterium extorquens* (score = 332 bits, identidade = 53%), *Beijerinckia indica* (score = 321 bits, identidade = 51%) e *Burkholderia phytofirmans* (score = 307 bits, identidade = 47%). Um aspecto interessante desta análise é que os organismos que tem catalases mais similares já foram todos isolados de folhas de plantas, têm capacidade de promover o crescimento de mudas de várias plantas e são capazes de fixar nitrogênio atmosférico [Abanda-Nkpwatt et al., 2006, Sessitsch et al., 2005, Spiff e Odu, 1973]. A sequência foi denominada *kat*.

A terceira sequência identificada (GDI2359) tem 718 resíduos. A busca por domínios conservados indicou que ela possui um domínio com atividade amidotransferásica "(GATase1)-like", este tipo de domínio é encontrado no C-terminal de várias catalases grandes. Este grupo inclui as grandes catalases de *Escherichia coli* HPII. Foi proposto que este domínio pode facilitar o processo de dobramento e oligomerização. A catalase HPII de *E. coli* é expressa na fase estacionária. A sequência analisada tem semelhança

com outras catalases de microorganismos fixadores de nitrogênio, como as catalases de *Rhodopseudomonas palustris* (score = 1085 bits, identidade = 74%), *Azoarcus sp.* (score = 1055 bits, identidade = 72%) e a *katC* de *Sinorhizobium meliloti* (score = 1019 bits, identidade = 70%). Denominamos esta *katC*.

As catalases GDI0467 e GDI2359 não possuem grande semelhança com organismos estreitamente relacionados, mas principalmente com outras bactérias endofíticas, um indicativo de que esses genes foram adquiridos mais tarde, possivelmente por transferência lateral. Isto poderia ser especialmente importante para a adaptação ao estilo de vida de bactérias endofíticas e pode conferir vantagens adaptativas a *G. diazotrophicus* em comparação com outros micróbios que colonizam o mesmo nicho.

Tabela 3.2: Tipos de catalase

	<i>E. coli</i>	<i>G. diazotrophicus</i>	<i>B. japonicus, R. etli</i>	<i>S. meliloti</i>	<i>M. loti</i>
HPII (mono-functional)	KatE	KatE(GDI0079) KatC(GDI2359)		KatA* KatC	KatE
HPI (bifunctional)	KatG*		KatG*	KatB	KatG*
Non-heme catalase		Kat (GDI0467)			

*Respondem à regulação pelo OxyR, induzido durante a fase exponencial e pela adição de H_2O_2

A anotação o genoma de *G. diazotrophicus* também indica a presença de dois genes para glutathione redutase (GDI2216 e GDI2280) no cromossomo principal. GDI2216 (GRA) também tem similaridade com os de *Alphaproteobacterias* intimamente relacionados, enquanto GDI2280 (GRB) tem semelhança com a glutathione redutase de organismos que vivem em associação ou simbiose com plantas, como *Stenotrophomonas sp.* (score = 446 bits, identidade = 57%), *Xanthomonas campestris* (score = 436 bits, identidade = 63%), e *Sinorhizobium meliloti* (score = 432 bits, identidade = 57%).

3.3.4 PCR em tempo real

Para determinar a contribuição individual de diferentes enzimas responsáveis pela detoxificação de ROS durante o processo de FBN, a expressão de seis genes – correspondentes as enzimas superóxido dismutase, catalase do tipo E, catalase, catalase do tipo C, glutathiona redutase tipo A e glutathiona redutase do tipo B – foi analisada por qPCR. Esses genes foram identificados no recém publicado genoma da bactéria [Bertalan et al, 2009] e foram escolhidos por serem diretamente envolvidos na detoxificação de ROS.

A expressão destes genes foi investigada em duas condições de crescimento: em meio líquido LGI-P suplementado com 1 ou 20 mM de sulfato de amônio. Considerando-se que as condições de FBN são aquelas em que a célula é crescida somente com 1 mM de sulfato de amônio no meio de cultura LGI-P (FIX) e as condições de não fixação aquelas em que a célula é crescida com 20 mM de sulfato de amônio no meio de cultura (NFIX).

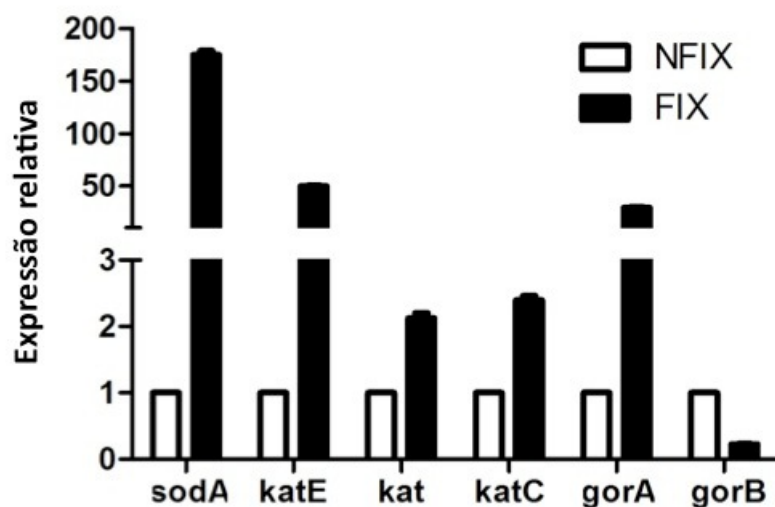


Figura 3.4: Expressão de enzimas de *G. diazotrophicus* envolvidas na detoxificação de ROS durante o processo de FBN. Expressão do mRNA correspondentes aos genes responsáveis pela detoxificação de ROS em células fixando e não-fixando nitrogênio crescidas durante 72h. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado pelo valor absoluto do gene na condição em que as células não fixam nitrogênio (NFIX). Os valores foram normalizados em relação a expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados.

Os resultados mostram que com 72h de crescimento, quase todos os genes analisados das células FIX apresentaram-se superexpressos quando comparados a condição NFIX, com exceção do *gorB*. Além do tempo de 72h a expressão gênica também foi analisada após 18, 24, 48 e 68 horas de crescimento. Em qualquer um dos tempos analisados os genes respondem de maneira similar a células crescidas por 72h. Uma exceção é a expressão do gene correspondente a enzima glutathione redutase B: ao contrario do resultado apresentado na Figura 3.4 apresentou-se superexpresso na condição FIX nos tempos de 18, 24, 48 e 68 horas de crescimento. Foi possível observar uma resposta generalizada da bactéria, aumentando a expressão dos genes de detoxificação de ROS durante o processo de FBN.

Esses dados sugerem que o padrão de expressão dos genes de detoxificação de ROS poderia ser correlacionado com os níveis de expressão de nitrogenase (*nifD*). Para avaliar essa hipótese, nós desenhamos oligos específicos capazes de amplificar o gene *nifD* da bactéria e repetimos o experimento de qPCR acrescentando a análise de expressão do gene *nifD* (Figura 4.9).

3.3.5 Agrupamento

Para identificar os genes de detoxificação de ROS mais relevantes para o processo de FBN foi realizada uma busca por padrões de expressão gênica entre os genes analisados. Os genes foram agrupados com base em seus padrões de expressão durante o crescimento e sob as condições de fixação e não-fixação de nitrogênio (como explicado no item anterior). Quanto mais próximos no gráfico de clusterização (Figura 3.5), mais similar a resposta entre eles.

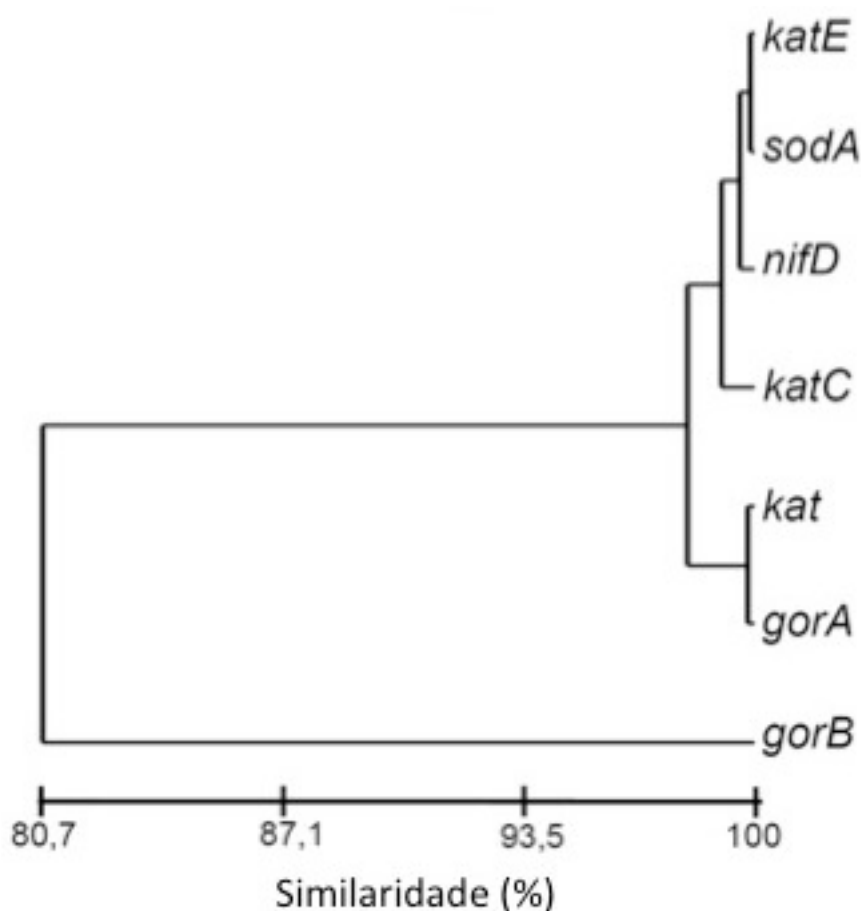


Figura 3.5: Construção de clusters de expressão dos genes de detoxificação de ROS revela a presença de dois clusters principais: um formado pela *gorB* sozinha e outro formado por *gorA*, *kat*, *katC*, *nifD*, *sodA* e *katE*. Esse cluster pode ser dividido em três clusters menores: *kat* + *gorA*, *katC* e *katE* + *sodA* + *nifD*.

A análise de clusters separou esses genes em dois clusters principais: um formado pela *gorB* sozinha e outro formado por *gorA*, *kat*, *katC*, *nifD*, *sodA* e *katE*. Esse segundo grupo apresenta já 98% de similaridade de resposta a presença ou ausência de nitrogênio suplementar no meio. Esse cluster pode ser dividido em três clusters menores: *kat* + *gorA*, *katC* e *katE* + *sodA* + *nifD*. O genes *katE* + *sodA* + *nifD* apresentam um padrão de resposta a presença de nitrogênio no meio igual (99,8% de similaridade de resposta), mesmo quando diferentes algoritmos são utilizados para a clusterização (single and average linkage, median linkage and ward linkage). Esse fato indica que as enzimas codificadas pelos genes *sodA* e *katE* devem ter uma importante papel protetor durante o processo de FBN.

3.3.6 Fixação de nitrogênio na presença de paraquat

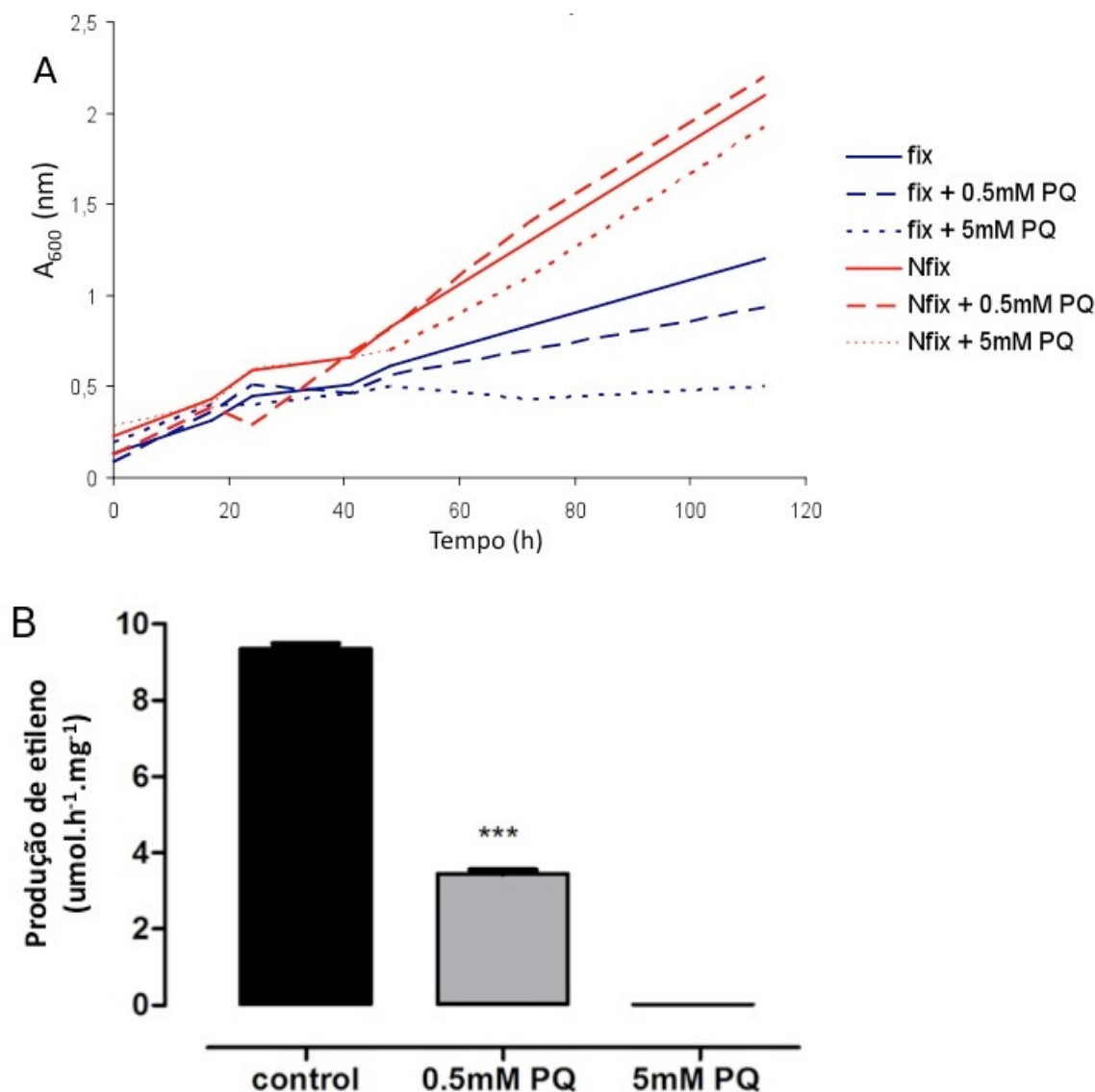


Figura 3.6: Efeitos de ROS no crescimento e na redução de acetileno ARA de células de *G. diazotrophicus*. A) Efeito de diferentes concentrações de paraquat (PQ) no crescimento de células fixando (em preto) e não-fixando (em cinza) nitrogênio. Depois de um período inicial de 24h, foi adicionado PQ para uma concentração final de 0,5 e 5 mM (linhas tracejadas e pontilhadas, respectivamente) (B) ARA de células depois de 72h de cultivo. A atividade foi medida depois de 48h na presença de 0,5 e 5mM de PQ.

A presença de ROS inibe fortemente o processo de FBN. Para explorar melhor as diferenças relativas aos níveis de ROS entre células crescidas em condição de fixação de nitrogênio ou não, a resistência da bactéria ao estresse oxidativo foi investigada pela adição de Paraquat (PQ) nas culturas. PQ é um potente gerador de ROS.

O crescimento das células fixando de nitrogênio foi fortemente inibido pela adição de PQ (Figura 3.6A). No entanto, células que não estavam fixando nitrogênio foram resistentes tanto a 0,5 quanto a 5 mM de PQ. Elas mantiveram a mesma taxa de crescimento e as mesmas densidade alcançadas (Figura 3.6A). A aparência geral das células foi similar para todas as condições testadas (não mostrado).

Para entender melhor as bases da inibição do crescimento das células fixadoras crescidas na presença de PQ, a atividade de nitrogenase foi avaliada pelo teste de redução do acetileno (ARA). As células crescidas sem adição de PQ foram usadas como controle. Como na curva de crescimento, PQ foi adicionado ao meio de cultivo 24h após o inóculo inicial e o teste de redução do acetileno 48h após a adição de PQ (72h de cultivo). A atividade de nitrogenase diminuiu 64% quando 0,5 mM de PQ foi adicionado ao meio de cultura. Na concentração de 5 mM de PQ, a atividade de nitrogenase foi completamente inibida. Esse resultado mostra que a presença de ROS abole o processo de FBN em *G. diazotrophicus* (Figura 3.6B).

3.4 Conclusões

Apesar de vários autores concordarem que a estratégia de proteção respiratória desempenha um papel importante na proteção da nitrogenase [Flores-Encarnacion et al., 1999, Pan e Vessey, 2001, Ureta e Nordlund, 2002] um aspecto que tem recebido pouca atenção é o impacto do aumento das taxas respiratórias na produção de ROS. Altas taxas respiratórias podem levar a um aumento na produção de ROS [Boveris e Chance, 1973], favorecendo a formação de um desequilíbrio redox e conseqüentemente a formação de estresse oxidativo. Nosso trabalho investigou o metabolismo redox durante a FBN e chegou às seguintes conclusões:

- Células crescidas em condições que a estimulem a fixar nitrogênio (FIX) apresentam uma diminuição do sinal de fluorescência (indicativo da presença de ROS) quando comparadas a células crescidas em condições de não-fixação.
- A presença de ROS em células não-fixando nitrogênio mostra que estas células habitam um ambiente onde ROS faz parte do meio celular sem exibir qualquer efeito adverso. Isso é consistente com a noção de que ROS agem como moléculas fisiológicas em um conjunto diversificado de situações biológicas [Forman et al., 2010, Terada, 2006], como parece ocorrer em células de *G. diazotrophicus* não fixando nitrogênio.
- A diminuição da quantidade de ROS intracelular durante o processo de FBN é resultado da superexpressão de enzimas detoxificadoras de ROS, especialmente a catalase do tipo E e a superóxido dismutase.
- A presença de altas concentrações de ROS intracelulares inibem a atividade de nitrogenase e conseqüentemente o crescimento de células com crescimento dependente de FBN.

Este estudo pode ser visto como o primeiro passo na compreensão do metabolismo redox em *Acetobacteraceae*. Aqui é apresentada a primeira caracterização da produção de ROS

e modificação do estado redox durante o processo de FBN. Nós propomos que durante as FBN, a *G. diazotrophicus* ative suas vias de detoxificação de ROS. Isso resulta numa diminuição dos níveis intracelulares de ROS e protege a nitrogenase da inibição por ROS. Esses resultados complementam a hipótese da proteção respiratória e permitem um melhor compreensão da atividade de nitrogenase em ambientes aeróbicos.

Capítulo 4

A via alternativa de aminoacilação de tRNA e o metabolismo de asparagina

4.1 Introdução

4.1.1 A síntese proteica

O dogma central da biologia molecular foi descrito em 1958 por Francis Crick na tentativa de relacionar o DNA, o RNA e as proteínas. Ele propõe que há uma unidirecionalidade na expressão da informação contida nos genes de uma célula, ou seja, o DNA é transcrito em RNA mensageiro e que este é traduzido em proteína. O dogma também postula que só o DNA pode replicar-se e, conseqüentemente, reproduzir-se e transmitir a informação genética para a descendência [Crick, 1970]. Ele pode ser resumido pela figura:



Figura 4.1: Dogma central da Biologia molecular proposto por Francis Crick em 1958

Apesar da descoberta dos vírus de RNA ter alterado ligeiramente o postulado neste dogma, ele ainda é uma estrutura que nos permite presumir de que maneira a informação contida na sequência de DNA finalmente se transforma em proteínas, que em última instância terão o papel de moléculas “efetoras” dentro do ambiente celular. Para realmente se compreender exatamente como ocorre a formação de proteínas é necessário um estudo mais detalhado do processo de tradução do código genético. Didaticamente, este processo pode ser dividido em cinco estágios:

1. Ativação dos aminoácidos: durante este estágio, cada um dos vinte e dois aminoácidos é covalentemente ligado ao tRNA cognato com gasto de ATP (adenosina trifosfato). Estas reações são catalisadas por uma classe de enzimas chamada de aminoacil-tRNA sintetases (aaRSs) (Figura 4.2) .

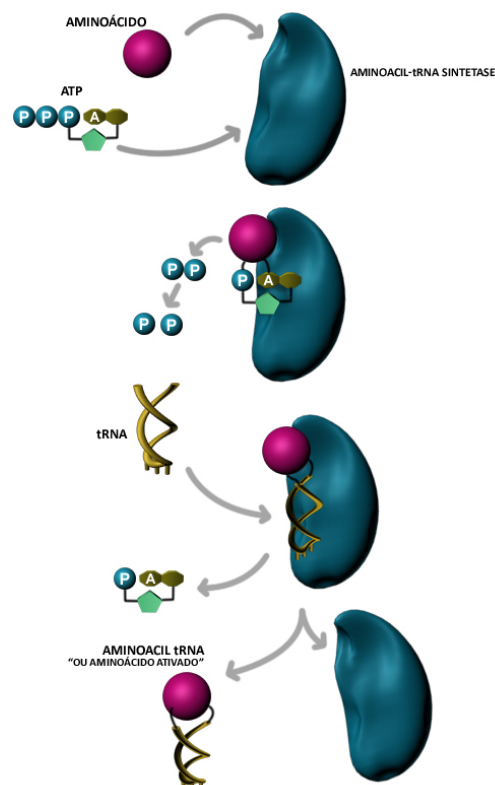


Figura 4.2: Mecanismo geral de atuação das aminoacil-tRNA sintetases. Primeiramente a enzima se liga ao aminoácido e o liga a uma molécula de AMP, proveniente da clivagem de um ATP. Em seguida a enzima liga esse aminoácido ao seu tRNA correspondente e libera a molécula de AMP.

2. Iniciação: o mRNA que codifica o polipeptídeo a ser sintetizado liga-se a subunidade menor do RNA ribossomal (rRNA) acompanhado da ligação do aminoacil-tRNA de iniciação e da subunidade maior do rRNA, formando o complexo de iniciação. As bases do tRNA aminoacilado pareiam-se com o códon AUG do mRNA, que sinaliza o início de uma cadeia polipeptídica. Este processo que requer GTP (guanosina trifosfato) é promovido por proteínas citossólicas chamadas de fatores de iniciação.
3. Alongamento: neste estágio a cadeia polipeptídica é alongada pela ligação covalente de unidades de aminoácidos sucessivamente, cada um levado ao ribossomo e corretamente posicionado pelo seu tRNA. O alongamento é promovido por proteínas citossólicas chamadas de fatores de alongamento. A ligação de cada aminoacil-tRNA e o movimento do ribossomo ao longo do mRNA ocorrem através da hidrólise de duas moléculas de GTP para cada resíduo adicionado a cadeia polipeptídica (esquema na Figura 4.3).

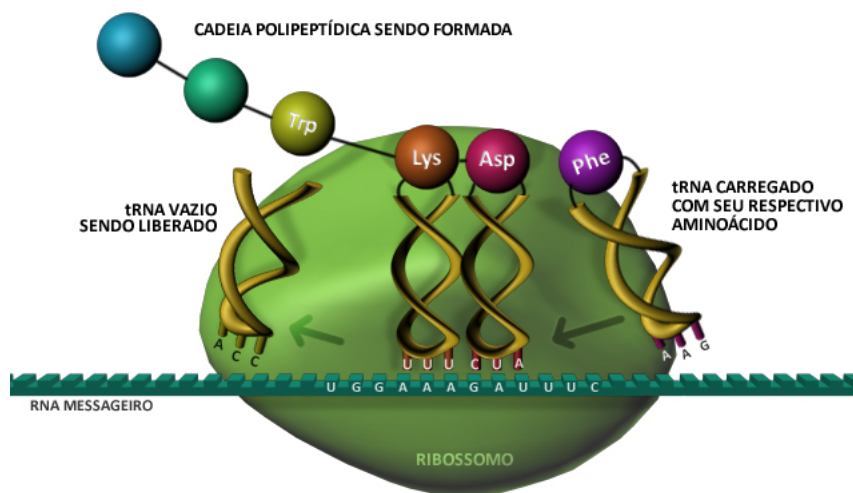


Figura 4.3: Mecanismo geral de síntese proteica.

4. Terminação e Liberação: a finalização da cadeia polipeptídica é sinalizada por um códon de terminação no mRNA (UAA, UAG ou UGA). A cadeia polipeptídica é então liberada do ribossomo por proteínas chamadas de fatores de liberação.

5. Processamento: o novo polipeptídeo pode sofrer processamento enzimático para remover um ou mais aminoácidos do terminal amino, para adicionar radicais acetil, fosfato, metil, carboxil ou ligar oligossacarídeos ou grupamentos prostéticos.

4.1.2 A via alternativa de aminoacilação de tRNA.

A formação do aminoacil-tRNA é um passo crucial para a síntese protéica. Durante este estágio, cada um dos vinte e dois aminoácidos é covalentemente ligado ao tRNA cognato com gasto de ATP (adenosina trifosfato). A ativação e ligação do aminoácido ao seu tRNA cognato são catalisadas pelas enzimas. A ligação do aminoácido ao tRNA é importante não apenas porque ativa a carboxila, mas também porque os aminoácidos por si mesmo não conseguem reconhecer os códons no mRNA. Por estes dois passos envolverem o reconhecimento específico do aminoácido e do tRNA, eles conferem especificidade a esta reação [McClain, 1993].

Seria racional assumir que todas as células possuiriam pelo menos vinte e duas aminoacil-tRNA sintetases, uma para cada aminoácido. Contudo, algumas observações derivadas de estudos bioquímicos e de genoma funcional de bactérias e principalmente de arqueas revelaram que uma grande parte dos organismos não utiliza todo o repertório de aaRS. Neste caso, na ausência de uma determinada aaRS, a síntese do aminoacil-tRNA é feita por uma rota evolutivamente antiga baseada na acilação errônea e subsequente modificação do aminoácido já acilado ao tRNA [Ambrogelly et al., 2007].

Em meados dos anos sessenta, foi descrita pela primeira vez a existência desta possível via alternativa para a síntese de $Gln - tRNA^{Gln}$ (Wilcox e Nirenberg, 1968). Neste trabalho foi demonstrado que a mesma enzima que acilava o glutamato no seu tRNA cognato ($tRNA^{Glu}$) também acilava o glutamato no tRNA da glutamina ($tRNA^{Gln}$). Sendo o produto $Gln - tRNA^{Gln}$ gerado somente após a conversão do $Glu - tRNA^{Gln}$ através de uma reação de transaminação dependente de ATP e na presença de um doador de grupo amina [Lapointe et al., 1986]. A enzima responsável pela conversão de $Glu - tRNA^{Gln}$ em $Gln - tRNA^{Gln}$ é chamada de Glutamina amidotransferase, ou Glu-AdT ou

AdT (Figura 4.4).

Em 1996, uma via semelhante para a formação de $Asn - tRNA^{Asn}$ foi descrita em *Haloferax volcanii* [Curnow et al., 1996]. Nesta arquea, uma acilação errônea pela AspRS (aspartil-tRNA sintetase) leva à formação de $Asp - tRNA^{Asn}$, em seguida, uma transamidação deste aspartato pela aspartil-tRNA amidotransferase (Asp-AdT) produz $Asn - tRNA^{Asn}$.

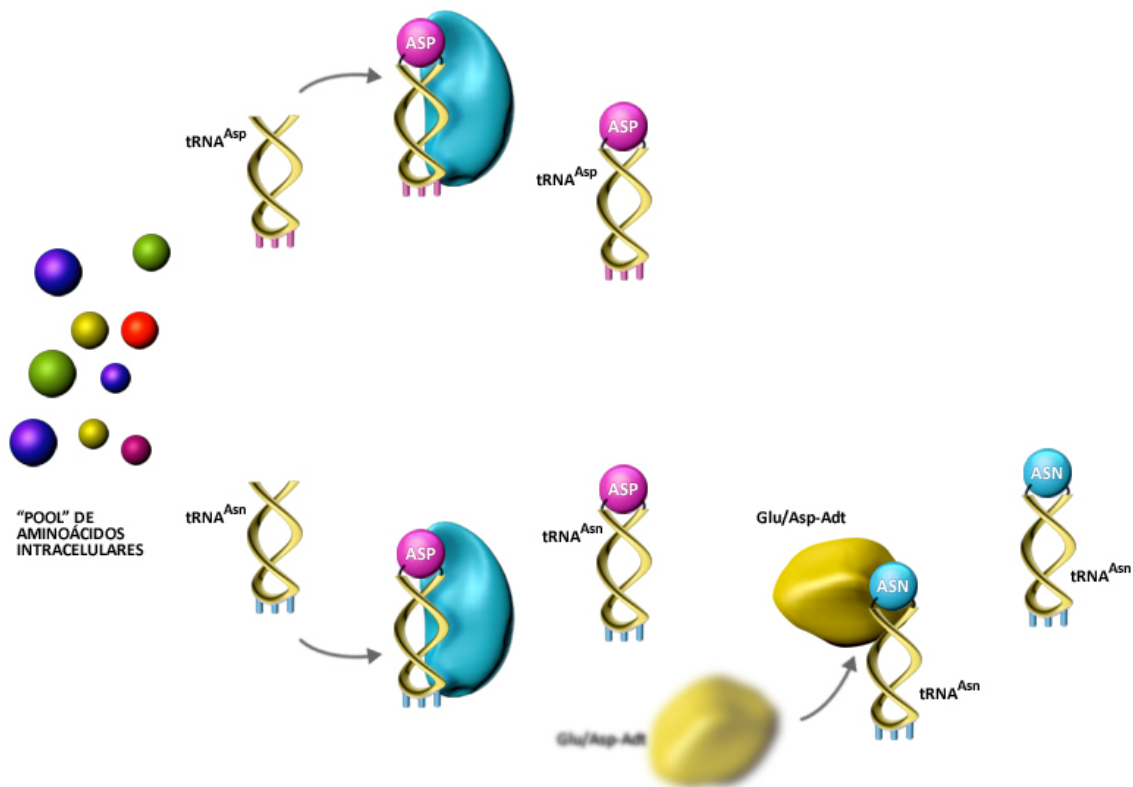
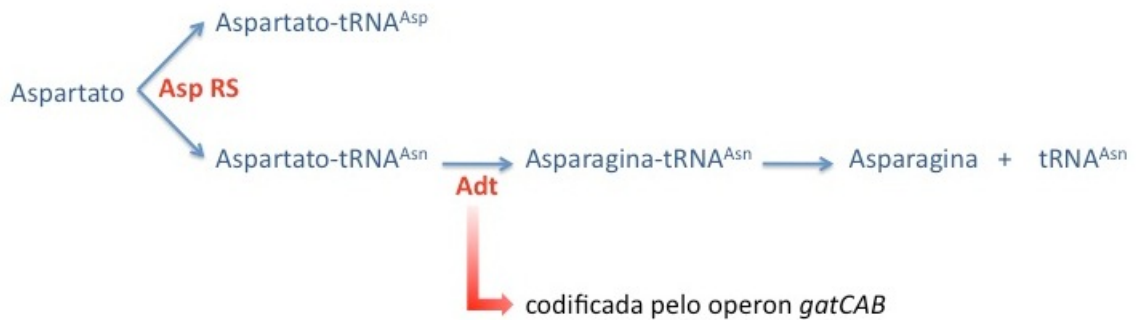


Figura 4.4: Esquema e cartoon da via para a formação de $Asn - tRNA^{Asn}$ (similar para $Gln - tRNA^{Gln}$). Uma AspRS não-discriminativa forma $Asp - tRNA^{Asp}$. O aminoacil-tRNA formado é então corrigido por uma amidotransferase (*gatCAB*) na presença de ATP e de um doador do grupamento amida.

A reação de transaminação realizada tanto para a transformação do glutamato em glutamina quanto do aspartato em asparagina é catalisada por uma única enzima glutamina/asparagina amidotransferase tRNA específica: Glu/Asp-AdT (Lapointe e colabora-

dores descreveram a Glu-Adt e Salazar e colaboradores mostraram que é uma enzima com afinidade tanto para Aspartato-tRNA quanto para Glutamato-tRNA) [Lapointe et al., 1986, Salazar et al., 2001]. Os genes codificantes para Glu/Asp-AdT se encontram em um complexo de três subunidades codificadas pelos genes *gatA*, *gatB*, e *gatC* [Curnow et al., 1997] ou pelo complexo de duas unidades codificadas pelos genes *gatD* e *gatE* em arqueas [Tumbula et al., 2000].

Genes homólogos ao *gatCAB* foram encontrados no genoma de todas as arqueas sequenciadas até hoje. A subunidade A possui grande homologia com uma classe de amidases e pode converter glutamina a glutamato e amônia in vitro. A subunidade B, está relacionada a uma classe de proteínas definidas com base na similaridade com o gene Pet 112 de levedura, que deve codificar uma subunidade da Glu-AdT mitocondrial de levedura, essencial para as funções da mitocôndria. A função da subunidade C não está totalmente clara, porém foi observada sua necessidade para a expressão da subunidade A [Curnow et al., 1997]. A atividade enzimática individual das subunidades revelou a ausência de atividade tanto das subunidades separadas quanto agrupadas duas a duas. A reação enzimática requer Mg^{+2} e ATP. Um olhar sobre sequências genômicas revela que as bactérias carregam duas formas de organização dos genes *gat*: algumas delas têm estruturas em operon, enquanto em outras os genes para as subunidades estão dispersos ao longo do cromossomo (por exemplo, arqueas e *Deinococcus radiodurans*).

A relevância da descoberta desta via alternativa tornou-se evidente quando foi mostrado que esta rota de acilação "errada", via transamidação poderia ser a via normal para a formação de *Gln* – *tRNA*^{*Gln*} em arqueas [White e Bayley, 1972], algumas bactérias Gram-positivas [Wilcox e Nirenberg, 1968, Wilcox, 1969, Schön et al., 1988b], cloroplastos [Schön et al., 1988a, Chen et al, 1990] e mitocôndrias [Schön et al., 1988a] além de duas bactérias Gram-negativas: *Rhizobium meliloti* [Gagnon et al., 1996] e *Helicobacter pylori* [Chuawong e Hendrickson, 2006]

A via indireta para formação de *Gln* – *tRNA*^{*Gln*} e *Asn* – *tRNA*^{*Asn*} existe quando a glutaminil-tRNS sintetase (GlnRS) ou a asparaginil-tRNA sintetase (AsnRS) estão ausen-

tes e uma glutamil-tRNA sintetase (GluRS) ou aspartil-tRNA sintetase (AspRS) promíscua, com dupla especificidade, está presente (Figura 4.4). Além disso, a existência de uma amidotransferase codificada pelo operon *gatCAB* é essencial para que o aminoacil-tRNA erroneamente carregado seja corrigido [Ibba et al., 2000].

4.1.3 A utilização da via alternativa para a produção de asparagina

Além da função essencial na síntese protéica, a enzima codificada pelo operon *gatCAB* recentemente foi descrita como uma enzima chave em outra via alternativa de alguns microrganismos: a via de formação de asparagina. O mecanismo de produção de asparagina através do *gatCAB* foi descrito na bactéria *Deinococcus radiodurans* por [Curnow et al., 1998].

A análise da sequência do genoma desta bactéria mostrou a presença de ambas as vias direta e indireta para a formação de Asn-tRNA and Gln-tRNA. Experimentos bioquímicos confirmaram essa observação mostrando a presença de asparaginil-tRNA sintetase (AsnRS) e glutaminil-tRNA sintetase (GlnRS), assim como uma GluRS e duas AspRS (uma promíscua e outra não promíscua). Além disso, tanto as atividades de transaminação sobre os substratos *Glu-tRNA^{Gln}* e *Asp-tRNA^{Asn}* foram verificadas. A presença de duas vias completas, aparentemente redundantes, para a formação de Asn-tRNA, sugeria que a o *gatCAB* poderia ter uma outra função no organismo.

Uma possível explicação para a presença incomum de vias redundantes de formação *Asn-tRNA^{Asn}* em *D. radiodurans* envolve um papel para o tRNA na biossíntese de aminoácidos. Até 1998 a única rota conhecida para a formação de asparagina era via asparagina sintetase, codificada pelos genes *asnA* ou *asnB* [Reitzer, 1996]. A asparagina sintetase dependente de glutamina, o produto do gene *asnB*, está presente em bactérias e eucariotos e utiliza ATP e glutamina para amidar um aspartato formando então uma molécula de asparagina [Humbert e Simoni, 1980]. A enzima codificada pelo gene *asnA*, é estruturalmente diferente da *asnB*, e usa como doador do grupo amida uma amônia em vez de glutamina [Cedar e Schwartz, 1968]. Em 1998, Curnow e colaboradores sugeriram

que a transaminação do aspartato ligado ao $tRNA^{Asn}$ poderia ser uma via ancestral de formação de asparagina em organismos que não possuem genes codificantes para asparagina sintetase e que esta poderia ser uma maneira de manter sua produção mais controlada [Curnow et al., 1998].

4.1.4 Influência de asparagina e glutamina na FBN

Para compreender a influência das condições nutricionais na atividade biológica de *G. diazotrophicus*, Tejera e colaboradores investigaram a atividade respiratória e da nitrogenase na presença de diferentes fontes de nitrogênio e carbono. Neste trabalho os autores mostram que a atividade da nitrogenase em *G. diazotrophicus* é afetada de acordo com a presença de aminoácidos como fonte de nitrogênio, verificando-se um efeito inibitório dependendo do aminoácido [Tejera et al., 2004] e da concentração de sacarose presentes no meio de cultivo [Reis e Döbereiner, 1998]. Asparagina, glutamina, aspartato, glutamato, cisteína e leucina inibem a atividade da nitrogenase entre 51% e 76% dependendo do aminoácido, sendo a asparagina, na concentração de 10 mM, o que apresentou maior efeito inibitório entre os aminoácidos testados (Tabela 4.1 - [Tejera et al., 2004]).

Guedes também mostrou resultados semelhantes em sua tese, mostrando que entre glutamina, glutamato, asparagina e aspartato a asparagina era o maior inibidor da atividade de nitrogenase (Tabela 4.1 - [Guedes, 2010])

Tabela 4.1: Influência de diferentes fontes de nitrogênio na atividade da nitrogenase, utilizando-se o método de redução do acetileno, em *G. diazotrophicus* crescida em meio LGI-P a 30°C após 72h de crescimento.

Concentração e tipo de fonte de nitrogênio usada para o crescimento	% de inibição da atividade de nitrogenase (por Tejera et al., 2004)	Concentração e tipo de fonte de nitrogênio usada para o crescimento	Atividade de nitrogenase (em nmoles de etileno/ μ g de proteína) (por Guedes, 2010)
1 mM NH ₄	0	-	-
5 mM NH ₄	18	-	-
15 mM NH ₄	32	-	-
10 mM Asn	76	2,5 mM Asn	75
10 mM Asp	68	2,5 mM Asp	90
10 mM Gln	67	2,5 mM Gln	110
10 mM Glu	60	2,5 mM Glu	125

4.2 Objetivos

O principal objetivo desta parte do trabalho foi esclarecer as vias de acilação de tRNA em *G. diazotrophicus*, o que podemos resumir com a pergunta:

1. existe uma via alternativa ativa de aminoacilação de tRNA neste organismo?

Alguns artigos sugerem que uma enzima presente nesta via alternativa de aminoacilação de tRNA é responsável pela síntese de asparagina pela célula [Curnow et al., 1998, Min et al., 2002]. Além disso tem sido mostrado que a regulação da enzima responsável pela fixação do nitrogênio atmosférico em amônia pode ser determinada principalmente pela oferta do aminoácido asparagina [Tejera et al., 2004, Guedes, 2010]. Desta forma, torna-se determinante o esclarecimento do processo de formação de asparagina e o envolvimento deste processo com o processo de fixação biológica de nitrogênio. Desta maneira, numa segunda etapa deste projeto procuramos responder as perguntas:

1. existe uma via alternativa de formação de asparagina ativa neste organismo?
2. a utilização de uma via alternativa de formação de asparagina pode influenciar a fixação de nitrogênio pela bactéria?
3. alterações no metabolismo de asparagina podem influenciar a atividade de nitrôgenase celular?

Foram traçados então os seguintes objetivos no intuito de responder as questões mencionadas acima:

- Identificação e análise dos genes das glutamil e aspartil-tRNA sintetases.
- Identificação e análise da Glutamil/Aspartil amidotransferase (AdT)
- Quantificação da expressão destes genes para verificar se a via alternativa de aminoacilação de tRNA está ativa e se responde à disponibilidade de nitrogênio.
- Busca por genes responsáveis pela síntese de asparagina.

- Verificar atividade da asparagina sintetase dependente de tRNA em *G. diazotrophicus*.
- Construção de mutantes aleatórios auxotróficos para asparagina.
- Análise das taxas de FBN dos mutantes.

4.3 Resultados e discussão

4.3.1 Localização e análise dos genes da via de aminoacilação de tRNA

A fim de investigar os mecanismos de aminoacilação de tRNA adotados por *G. diazotrophicus*, foram identificados no genoma recentemente publicado da bactéria [Bertalan et al, 2009] os genes responsáveis pelo processo de acilação de tRNA. A anotação do genoma indica a presença dos seguintes genes:

Aminoacil tRNA sintetases:

1. Ácido aspártico (Asp) - Aspartyl-tRNA synthetase (*aspS*, GDI2232)
2. Ácido glutâmico (Glu) -
 - (a) Glutamyl-tRNA synthetase I (*gltX*, GDI0050)
 - (b) Glutamyl-tRNA synthetase (*gltX*, GDI2211)
3. Alanina (Ala) - Alanyl-tRNA synthetase (*alaS*, GDI2187)
4. Arginina (Arg) - Arginyl-tRNA synthetase (*argS*, GDI2225)
5. Asparagina (Asn) - **Não encontrada**
6. Cisteína (Cys) - Cysteinyl-tRNA synthetase (*cysS*, GDI2199)
7. Fenilalanina (Phe) - Phenylalanyl-tRNA synthetase beta chain (*pheS*, GDI3271); Phenylalanyl-tRNA synthetase alpha chain (GDI3272)
8. Glicina (Gly) - Glycyl-tRNA synthetase alpha chain (*glyQ*, GDI2331); Glycyl-trna synthetase beta chain (*glyS*, GDI2332)
9. Glutamina (Gln) - **Não encontrada**
10. Histidina (His) - putative Histidyl-tRNA synthetase (GDI2256)
11. Isoleucina (Ile) - putative Isoleucyl-tRNA synthetase (*ileS*, GDI0589)

12. Leucina (Leu) - putative Leucyl-tRNA synthetase (*leuS*, GDI2393)
13. Lisina (Lys) -
 - (a) putative Lysyl-tRNA synthetase (*poxA*, GDI0509)
 - (b) putative Lysyl-tRNA synthetase (*lysS*, GDI2046)
14. Metionina (Met) - putative Methionyl-tRNA synthetase (GDI3217)
15. Pirrolisina - **Não encontrada**
16. Prolina (Pro) -
 - (a) putative prolyl-tRNA synthetases (GDI0897)
 - (b) Prolyl-tRNA synthetase (*proS*, GDI3024)
17. Selenocisteína - **Não encontrada**
18. Serina (Ser) - Seryl-tRNA synthetase (*serS*, GDI2219)
19. Tirosina (Tyr) - putative tyrosyl-tRNA synthetase (*tyrS*, GDI2375)
20. Treonina (Thr) - Threonyl-tRNA synthetase (*thrS*, GDI1448)
21. Triptofano (Trp) - putative Tryptophanyl-tRNA synthetase (*trpS*, GDI1855)
22. Valina (Val) - Valyl-tRNA synthetase (*valS*, GDI1763)

As enzimas aminoacil-tRNAs sintetases referentes a cada um dos aminoácidos existentes foram encontradas no genoma de *G. diazotrophicus*, com exceção das aa-tRNA sintetases dos aminoácidos asparagina, glutamina, pirrolisina e selenocisteína. A ausência e aa-tRNAs sintetases específicas para asparagina e glutamina sugere a existência de uma possível via alternativa para a síntese de *Asn - tRNA^{Asn}* e *Gln - tRNA^{Gln}*. Para que essa via alternativa seja efetiva, é necessário que:

1. as enzimas aspartil-tRNA sintetases e glutamil-tRNA sintetases sejam promíscuas, ou seja, sejam capazes de acilar o aspartato tanto no tRNA do aspartato ($tRNA^{Asp}$) quanto no tRNA da asparagina ($tRNA^{Asn}$) e o glutamato no tRNA do glutamato ($tRNA^{Glu}$) e da glutamina ($tRNA^{Gln}$).
2. exista uma enzima amidotransferase codificada pelos genes *gatA*, *gatB* e *gatC* (*gat-CAB*), que transfira um grupo amida para o aspartato ou glutamato acilado erroneamente.

Análise da sequência da Aspartil-tRNA sintetase (*aspS*, GDI2232)

O tRNA acilado erroneamente é sintetizado por uma aspartil-tRNA sintetase não discriminativa (ND-AspRS) que acila o aspartato tanto no $tRNA^{Asp}$ quanto no $tRNA^{Asn}$. Em contrapartida, uma aspartil-tRNA sintetase discriminativa (D-AspRS) forma somente $Asp - tRNA^{Asp}$. Foi realizada então uma análise da sequência da aspartil-tRNA sintetase encontrada no genoma de *G. diazotrophicus* buscando-se identificar características que a confirmassem como uma AspRS promíscua. Até 2005 não era possível identificar a especificidade de uma AspRS baseando-se somente em sua sequência de aminoácidos, mas o trabalho de Feng e colaboradores, de 2005, identificou a existência de uma prolina na posição 77 da alça L1 de AspRSs não discriminativas de arqueas que é responsável pela propriedade não discriminativa desta enzima. Substituições deste aminoácido por cisteína, isoleucina, leucina, lisina, fenilalanina, serina, ou valina mostram que os mutantes continuavam sendo capazes de formar $Asp - tRNA^{Asp}$ mas tornavam-se incapazes de reconhecer o $tRNA^{Asn}$ e conseqüentemente formar $Asp - tRNA^{Asn}$ [Feng et al., 2005].

É interessante notar que a alça correspondente à L1 em AspRSs bacterianas são diferentes na sequência e comprimento e não apresentam essa conservação tão forte da prolina. Conseqüentemente um outro mecanismo, ainda não completamente esclarecido deve ser adotado por essa classe de microrganismos para a distinção entre ND-AspRS e D-AspRS.

Experimentos com alinhamentos de sequências bacterianas e mutações dirigidas che-

garam no entanto a alguns sítios que parecem ser determinantes para o reconhecimento do $tRNA^{Asn}$ em *Helicobacter pylori*. Eles são as Leucinas 81 e 86, correspondentes à Asparagina 82 e Metionina 87 de *Escherichia coli*. Mutações nestes resíduos diminuíram a capacidade da enzima reconhecer o $tRNA^{Asn}$ e aumentaram a sua especificidade para $tRNA^{Asp}$ [Chuawong e Hendrickson, 2006].

Nós realizamos uma análise da sequência que codifica para a AspRS de *G. diazotrophicus* (Figura 4.5) e pudemos observar a existência de duas características que indicam que a AspRS desta bactéria poderia reconhecer o $tRNA^{Asn}$. Elas são:

- A presença de uma treonina (T) na posição 98 da sequência. Em AspRS discriminativas o aminoácido encontrado nesta posição é a asparagina (N). Em AspRS promíscuas é verificada a presença de principalmente treonina (T), leucina (L) ou alanina (A).
- A presença de uma leucina (L) na posição 103 da sequência. Em *H. pylori* a presença de uma L nesta posição diminui a especificidade por $tRNA^{Asp}$ e aumentaram a sua especificidade para $tRNA^{Asn}$, o que significa que a enzima se tornaria mais promíscua.

```

[Bpertussis      LAQIVFDPDPAEAFGTAERLRNEFCISITGLVRLRPEGTANAELASGEVEVLCQQVEILN 103
[Ecoli           IVQVFFDPDRADALKLASELRNEFCIQVTGTVRARDEKNINRDMATGEIEVLASSLTIIN 103
[Hpylori         LVQLVCDP-SSKAYEKALEVRSEFVLVAKGKVLRGAGLENPKLKTGKIEIVLEELIEN 102
[Gdiazotrophicus MTQIVIPA-GSPVLEAVERVRVESVLTITGEVLRDAATKPNLPTGEIELRAQQVDIQS 119
[Bmelitensis     ITQIVADP-DSPAFKVAETVRGEWVIRVDGEVKARADDAVNTNLPTGEVEIFATEIEVLS 105
::*:. . : . . : * * : * * * * * . *::* . : : .

```

Figura 4.5: Alinhamento da porção N-terminal de algumas aspartyl-tRNA sintetases bacterianas não discriminativas. A porção azul, indica a posição correspondente à Prolina 77 de arqueas. Em bactérias, essa é uma posição conservada, onde encontra-se sempre uma arginina, independente de ser uma aa-tRNS sintetase discriminativa ou não. A posição vermelha, correspondente à asparagina (N) de *E. coli*, em enzimas não discriminativas encontramos sempre Treonina (T), Leucina (L) ou Alanina (A). Já na porção marcada de amarelo (M87 de *E. coli*) a presença de uma leucina (L) diminui fortemente a especificidade por $tRNA^{Asp}$ e aumenta para $tRNA^{Asn}$. Foram usadas as seqüências dos seguintes organismos: *Bordetella pertussis* Tohama I, *Escherichia coli* str. K-12 substr. W3110, *Helicobacter pylori* 26695, *Gluconacetobacter diazotrophicus* PAI 5, *Brucella melitensis* bv. 1 str. 16M.

gatCAB

A procura pelos genes responsáveis pela transamidação do aminoácido acilado erroneamente confirmou a presença destes organizados num operon *gatCAB* no genoma da bactéria. A presença destas proteínas é comum em todas as arqueas [White e Bayley, 1972], em muitas bactérias Gram-positivas [Wilcox, 1969, Schön et al., 1988b], mas foi somente caracterizada em duas bactérias Gram-negativas: *Rhizobium meliloti* [Gagnon et al., 1996] e *Helicobacter pylori* [Chuawong e Hendrickson, 2006].

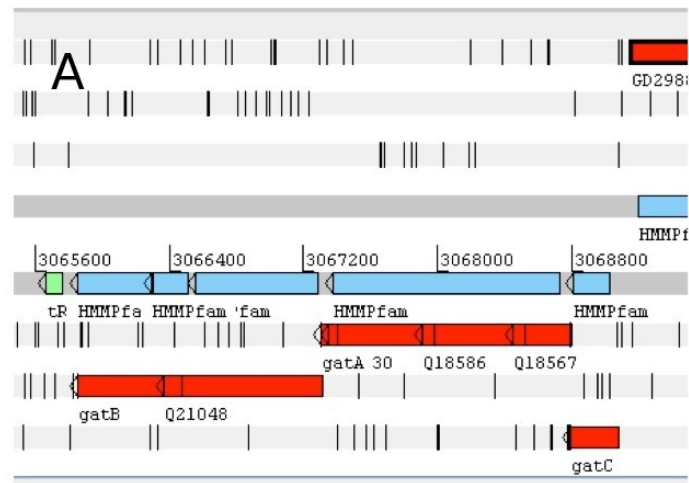


Figura 4.6: Localização do *gatCAB* no genoma da *G. diazotrophicus*.

4.3.2 Análise da expressão dos genes da via alternativa de síntese protéica

Para confirmar a utilização pela bactéria dos genes envolvidos na via alternativa de acilação do tRNA foi realizado um experimento de qPCR onde acompanhou-se a expressão destes genes em função da disponibilidade de nitrogênio. Para isso células de *G. diazotrophicus* foram crescidas em meio líquido LGI-P suplementado com 1 ou 20 mM de sulfato de amônio. Foi estabelecido que as condições de fixação são aquelas em que a célula é crescida somente com 1 mM de sulfato de amônio no meio de cultura e as condições de não fixação aquelas em que a célula é crescida com 20 mM de sulfato de amônio

no meio de cultura. Após 72h de crescimento realizou-se a extração do RNA e síntese das primeiras fitas de cDNA que foram utilizadas nas reações do qPCR para analisar a expressão dos genes *gluS*, *aspS*, *gatA*, *gatB* e *gatC*. Os resultados são apresentados na Figura 4.7

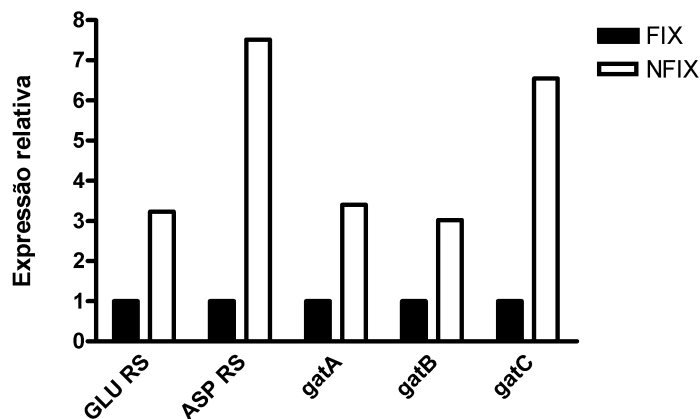


Figura 4.7: Expressão de enzimas da via alternativa de aminoacilação de tRNA em função da disponibilidade de nitrogênio - células crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIx) de sulfato de amônio - após 72h de crescimento. A expressão dos genes na condição de FIX de crescimento foi utilizada como calibradora e, portanto, fixada em 1. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado (Ct) pelo valor absoluto do gene na condição calibradora. Os valores foram normalizados em relação à expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados.

Os resultados mostram que, em células crescidas com 20 mM de sulfato de amônio, todos os genes responsáveis pela da via alternativa de acilação do tRNA tem sua expressão aumentada quando comparados a células crescidas com 1 mM de sulfato de amônio no meio. A expressão gênica também foi analisada após 18, 24, 36 e 48 horas de crescimento e os genes responderam de forma semelhante ao experimento realizado com 72h de crescimento (dados não mostrados).

Esses dados mostram que os genes responsáveis pela da via alternativa de acilação do tRNA são expressos e regulados em função da disponibilidade de nitrogênio orgânico. Observações em *Rhizobium* e *Anabaena* mostram que durante o processo de FBN a síntese de proteínas é inibida em relação a células que não estão fixando nitrogênio

[Ludwig, 1980, Wealand et al., 1989]. O resultado obtido no nosso experimento mostra que aparentemente o mesmo ocorre em *G. diazotrophicus*: os genes da via alternativa de acilação de tRNA são transcritos em condições que favorecem a síntese proteica.

4.3.3 Influência de diferentes fontes de nitrogênio na expressão do gene da nitrogenase em função do tempo de crescimento

Tejera e colaboradores (2004) e Guedes (2010) mostraram que a atividade de nitrogenase em *G. diazotrophicus* é profundamente afetada de acordo com a fonte de nitrogênio fornecida [Tejera et al., 2004, Guedes, 2010]. Desta maneira, decidimos investigar a influência dos aminoácidos envolvidos na via alternativa de acilação do tRNA na expressão da subunidade Mo-Fe da nitrogenase (gene *nifD*). Para isso células de *G. diazotrophicus* foram crecidas em meio LGI-P suplementado com 1 ou 20mM de sulfato de amônio ou suplementado com 20mM de asparagina, aspartato, glutamina ou glutamato. As curvas de crescimento são apresentadas na Figura 4.8 e podemos concluir a partir delas que não há alterações marcantes nas curvas de crescimento das células em nenhuma das condições testadas.

Curva de crescimento de *G. diazotrophicus* em meio suplementado com diferentes fontes de nitrogênio

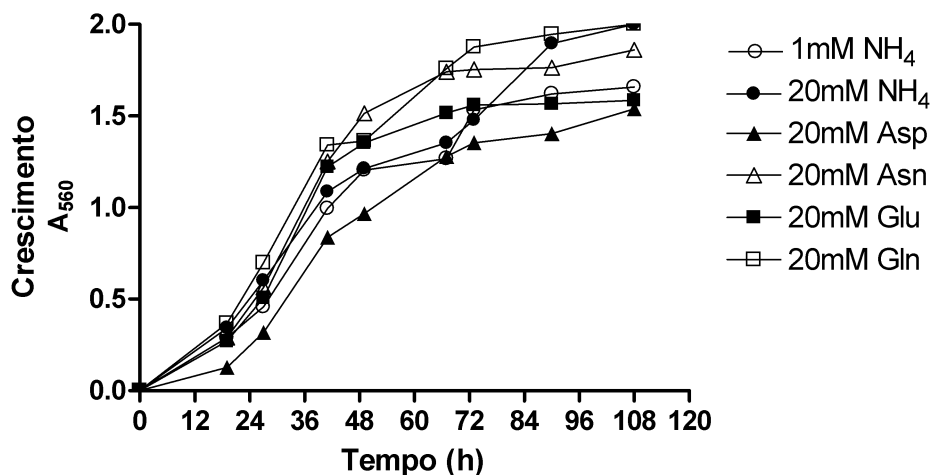


Figura 4.8: Crescimento de *G. diazotrophicus* em meio LGI-P líquido suplementado com 1 ou 20mM de sulfato de amônio ou suplementado com 20mM de asparagina, aspartato, glutamina ou glutamato. As células foram cultivadas a 30°C, 150 rpm.

Influência de diferentes fontes de nitrogênio na expressão do gene da nitrôgenase

A influência das diferentes fontes de nitrogênio na expressão do gene *nifD* foi analisada através de um experimento de qPCR onde acompanhou-se a expressão deste gene em diferentes tempos de cultivo (25, 72 e 96h) em células crescidas com diferentes aminoácidos como doadores de nitrogênio. Os resultados são apresentados na Tabela 4.2.

Tabela 4.2: Influência de diferentes fontes de nitrogênio na expressão do gene da nitrogenase em função do tempo de crescimento

Concentração e tipo de fonte de nitrogênio usada para o crescimento	% de inibição da expressão do gene da nitrogenase (25h)	% de inibição da expressão do gene da nitrogenase (72h)	% de inibição da expressão do gene da nitrogenase (96h)
1 mM NH ₄	0	0	0
20 mM NH ₄	60	89	96
20 mM Asn	62	92	97
20 mM Asp	58	81	88
20 mM Gln	32	79	86
20 mM Glu	49	84	91

Em todas as condições analisadas a asparagina é o maior inibidor da expressão do gene *nifD*. Este resultado é especialmente interessante sob a ótica de um outro trabalho, onde os autores mostram que a asparagina é o principal aminoácido na cana-de-açúcar, representando de 72-85% dos aminoácidos da planta [Loiret et al., 2009].

Considerando que o apoplasto é o principal nicho para organismos endofíticos [McCully, 2001] que consideráveis quantidades de asparagina foram encontradas no espaço apoplástico da cana-de-açúcar [Tejera et al., 2006, Loiret et al., 2009], e que ela é o aminoácido que mais influencia a expressão (este trabalho) e atividade da nitrogenase de *G. diazotrophicus* [Tejera et al., 2006, Guedes, 2010] pode-se dizer que essa forma de nitrogênio possui um papel fundamental na regulação do processo de FBN em *G. diazotrophicus*.

Quantificação de aminoácidos livres no interior da célula

Embora aparentemente a asparagina seja um aminoácido com capacidade inibitória da nitrogenase muito alta, geralmente seus níveis intracelulares são mantidos muito baixos [Capone et al., 1994, Hua et al., 1982]. Além disso, altas concentrações extracelulares de asparagina não significam necessariamente altas concentrações intracelulares do aminoácido, pois seu transporte é mediado por proteínas específicas. Desta maneira, a inibição da expressão e atividade da nitrogenase observadas quando a asparagina é adicionada ao meio

de crescimento não significa necessariamente que ele é o principal inibidor intracelular.

Para investigar os efeitos da asparagina intracelularmente foi realizada uma dosagem das concentrações intracelulares dos aminoácidos asparagina, glutamina, glutamato e aspartato utilizando-se cromatografia líquida de alto desempenho (HPLC). Os resultados são apresentados na tabela 4.3.

Tabela 4.3: Concentração em μM medida por HPLC de aminoácidos livres intracelulares de células de *G. diazotrophicus* crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIX) de NH_4SO_4 em função do tempo de cultivo.

TEMPO (h)	Asparagina (μM)		Glutamina (μM)		Aspartato (μM)	
	FIX	NFIX	FIX	NFIX	FIX	NFIX
24	40	59	144	259	3,3	3,4
48	38	66	124	268	3,3	4,5
68	33	72	152	191	3,8	3,8
72	31	73	338	161	4,0	4,4

Relação entre a quantidade de asparagina com a expressão do gene da nitrógenase

Para determinar a influência das concentrações intracelulares dos aminoácidos asparagina, aspartato e glutamina no processo de FBN, a expressão do gene *nifD* foi analisada por qPCR e relacionada com as quantidades intracelulares dos aminoácidos testados. A expressão do *nifD* foi investigada nas mesmas condições e tempos de crescimento do experimento anterior, em que dosamos as quantidades intracelulares dos aminoácidos. Isto é, com 1 mM de sulfato de amônio como fonte de nitrogênio no meio de cultura (condição que estimula o processo de FBN - FIX) e com 20 mM de sulfato de amônio no meio (condição que reprime o processo de FBN - NFIX). A expressão do gene da *nifD* ao longo do tempo nas condições testadas pode ser observada na Figura 4.9.

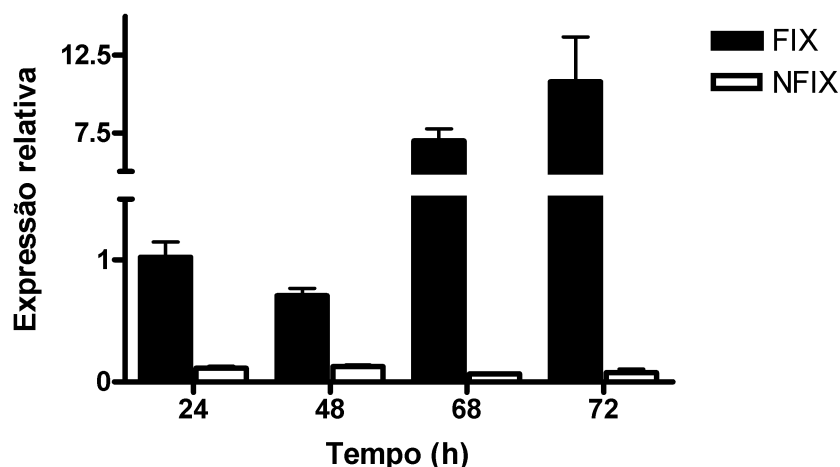


Figura 4.9: Expressão do gene *nifD* em função da disponibilidade de N no meio de cultura. As células foram crescidas em meio LGI-P suplementado com 1 (FIX) ou 20 mM (NFIX) de sulfato de amônio - e a expressão relativa do gene *nifD* foi verificada. A expressão do gene na condição FIX com 24h de cultivo foi utilizada como calibradora e, portanto, fixada em 1. O nível dos transcritos é representado como a taxa (expressão relativa) do valor absoluto do gene estudado (Ct) pelo valor absoluto do gene na condição calibradora. Os valores foram normalizados em relação à expressão do gene constitutivo 23S. Média +/- barra de erros de três experimentos independentes são mostrados.

Os resultados mostram que o gene *nifD* de células crescidas em condições de fixação de nitrogênio está sempre superexpresso quando comparado à condição que reprime o processo de FBN. A busca por uma relação entre os níveis intracelulares dos aminoácidos analisados e a expressão do *nifD* (Figura 4.10) mostrou que a asparagina é o único aminoácido que pode ter seus níveis correlacionados com a expressão do *nifD*. Esse resultado mostra que além da asparagina ser o principal inibidor do processo de FBN e da expressão do *nifD* quando ela está presente no meio extracelular, ela parece ser também o principal inibidor intracelular, entre os aminoácidos testados, da expressão do gene *nifD* e conseqüentemente do processo de FBN.

também está presente com a finalidade formar asparagina neste organismo.

Nós realizamos então uma busca em 150 genomas completamente sequenciados naquele momento por genes codificantes para a asparagina sintetase do tipo A ou B. Nossa análise revelou que todos os organismos que não possuíam genes ortólogos a *asnA* ou *asnB*, mesmo que apresentassem o gene *asnS* (que codifica para a asparaginil-tRNA sintetase), possuíam os genes *gatCAB* (que codificam para a Asp/Glu-tRNA amidotransferase) (Tabela 4.4). Isso é uma indicação de que nestes organismos a síntese de asparagina deve estar condicionada a ação da enzima AdT, codificada pelo *gatCAB*, como seria também o caso de *G. diazotrophicus*

Tabela 4.4: Identificação da enzima gatCAB, da via alternativa de acilação de tRNA, em todos os organismos que não possuem os genes que codificam para as enzimas asparagina sintetase A ou B (*asnA* e *asnB*).

Gênero	<i>asnA</i>	<i>asnB</i>	<i>gatCAB</i>	<i>asnS</i>	<i>NDaspS</i>	<i>glnS</i>
<i>Acidobacteria</i>	1	0	1	1	0	0
<i>Acinetobacter</i>	0	0	1	0	1	1
<i>Agrobacterium</i>	0	0	1	0	1	0
<i>Alcanivorax</i>	0	0	1	0	1	1
<i>Anabaena</i>	0	0	1	1	0	0
<i>Anaeromyxobacter</i>	0	1	1	1	0	1
<i>Anaplasma</i>	0	0	1	0	1	0
<i>Aquifex</i>	0	0	1	0	1	0
<i>Aster yellows</i>	0	1	0	1	0	1
<i>Azoarcus</i>	0	1	1	0	1	1
<i>Bacillus</i>	1	1	1	1	0	0
<i>Bacillus</i>	0	1	1	1	0	0
<i>Bacteroides</i>	1	1	0	1	0	1
<i>Bartonella</i>	0	0	1	0	1	0
<i>Baumannia</i>	0	0	0	1	0	1
<i>Bdellovibrio</i>	0	0	1	1	0	1
<i>Bifidobacterium</i>	0	0	1	0	1	0
<i>Bordetella</i>	0	1	1	0	1	1
<i>Borrelia</i>	0	0	1	1	0	0
<i>Bradyrhizobium</i>	0	1	1	0	1	1
<i>Brucella</i>	0	0	1	0	1	0
<i>Buchnera</i>	0	0	0	1	0	1
<i>Burkholderia</i>	0	1	1	0	1	1
<i>Campylobacter</i>	0	0	1	0	1	0
<i>Candidatus Blochmannia</i>	0	0	0	1	0	1
<i>Candidatus Pelagibacter</i>	0	0	1	0	1	0
<i>Candidatus Protochlamydia</i>	0	0	1	1	0	0
<i>Carboxydotherrmus</i>	0	1	1	0	1	0
<i>Caulobacter</i>	0	0	1	0	1	0
<i>Chlamydia</i>	0	0	1	0	1	0
<i>Chlamydothila</i>	0	0	1	0	1	0
<i>Chlorobium</i>	0	0	1	0	1	0
<i>Clostridium</i>	0	1	1	1	1	1
<i>Clostridium</i>	1	0	0	1	0	1
<i>Corynebacterium</i>	0	1	1	0	1	0
<i>Coxiella</i>	0	1	1	0	1	0
<i>Cytophaga</i>	0	1	1	1	0	0
<i>Dechloromonas</i>	0	1	1	0	1	1
<i>Dehalococcoides</i>	0	1	1	0	1	0
<i>Deinococcus</i>	0	0	1	1	1	1
<i>Desulfitobacterium</i>	0	1	1	1	0	1
<i>Desulfotalea</i>	0	1	1	1	0	1
<i>Desulfovibrio</i>	0	1	1	1	0	1
<i>Ehrlichia</i>	0	0	1	0	1	0
<i>Enterococcus</i>	0	0	1	1	0	0
<i>Erwinia</i>	1	1	0	1	0	1
<i>Erythrobacter</i>	0	1	1	0	1	0
<i>Escherichia</i>	1	1	0	1	0	1
<i>Francisella</i>	0	1	1	0	1	1
<i>Frankia</i>	0	1	1	0	1	0
<i>Fusobacterium</i>	1	0	1	1	0	0

<i>Geobacillus</i>	0	1	1	1	0	0
<i>Geobacter</i>	0	1	0	1	0	1
<i>Gluconoacetobacter</i>	0	0	1	0	1	0
<i>Gloeobacter</i>	0	0	1	0	1	0
<i>Gluconobacter</i>	0	1	1	0	1	0
<i>Haemophilus</i>	1	0	0	1	0	1
<i>Hahella</i>	0	1	1	1	0	1
<i>Helicobacter</i>	0	0	1	0	1	0
<i>Idiomarina</i>	1	1	0	1	0	1
<i>Jannaschia</i>	0	0	1	0	1	0
<i>Lactobacillus</i>	1	1	1	1	0	0
<i>Lactococcus</i>	0	1	1	1	0	0
<i>Lawsonia</i>	0	1	1	1	0	1
<i>Legionella</i>	0	0	1	1	0	1
<i>Legionella</i>	0	0	1	1	0	0
<i>Leifsonia</i>	0	1	1	0	1	0
<i>Leptospira</i>	0	1	1	1	0	0
<i>Listeria</i>	0	1	1	1	0	0
<i>Magnetospirillum</i>	0	1	1	0	1	0
<i>Mannheimia</i>	1	0	0	1	0	1
<i>Mesoplasma</i>	0	0	1	1	0	0
<i>Mesorhizobium</i>	0	1	1	0	1	0
<i>Methylobacillus</i>	0	0	1	0	1	1
<i>Methylococcus</i>	0	1	1	0	1	0
<i>Moorella</i>	0	1	1	0	1	0
<i>Mycobacterium</i>	0	1	1	0	1	0
<i>Mycoplasma</i>	1	0	1	1	0	0
<i>Mycoplasma</i>	0	0	1	1	0	0
<i>Myxococcus</i>	0	1	1	1	0	1
<i>Neisseria</i>	0	0	1	0	1	1
<i>Neorickettsia</i>	0	0	1	0	1	0
<i>Nitrobacter</i>	0	0	1	0	1	1
<i>Nitrosococcus</i>	0	1	1	0	1	0
<i>Nitrospira</i>	0	1	1	0	1	1
<i>Nocardia</i>	0	1	1	0	1	0
<i>Nostoc</i>	0	0	1	1	0	0
<i>Novosphingobium</i>	0	1	1	0	1	0
<i>Oceanobacillus</i>	0	1	1	1	0	0
<i>Onion yellows</i>	0	1	0	1	0	1
<i>Pasteurella</i>	1	0	0	1	0	1
<i>Pelobacter</i>	0	0	1	1	0	1
<i>Pelodictyon</i>	0	1	1	0	1	0
<i>Photobacterium</i>	0	1	0	1	0	1
<i>Polaromonas</i>	0	0	1	0	1	1
<i>Porphyromonas</i>	0	0	0	1	0	1
<i>Prochlorococcus</i>	0	1	1	0	1	0
<i>Pseudoalteromonas</i>	0	1	0	1	0	1
<i>Pseudoalteromonas</i>	1	1	0	1	0	1
<i>Pseudomonas</i>	0	1	1	0	1	1
<i>Psychrobacter</i>	0	1	1	0	1	1
<i>Ralstonia</i>	0	0	1	0	1	1
<i>Rhizobium</i>	0	1	1	0	1	0
<i>Rhodobacter</i>	0	0	1	0	1	0

<i>Rhodoferax</i>	0	1	1	0	1	1
<i>Rhodopirellula</i>	0	1	1	1	0	1
<i>Rhodopseudomonas</i>	0	1	1	0	1	1
<i>Rhodospirillum</i>	0	1	1	0	1	0
<i>Rickettsia</i>	0	0	1	0	1	0
<i>Roseobacter</i>	0	1	1	0	1	0
<i>Rubrobacter</i>	0	0	1	0	1	0
<i>Saccharophagus</i>	0	1	1	0	1	1
<i>Salinibacter</i>	0	0	1	1	0	1
<i>Salmonella</i>	1	1	0	1	0	1
<i>Shewanella</i>	0	1	0	1	0	1
<i>Shigella</i>	1	1	0	1	0	1
<i>Silicibacter</i>	0	0	1	0	1	0
<i>Sinorhizobium</i>	0	1	1	0	1	0
<i>Sodalis</i>	0	1	0	1	0	1
<i>Sphingopyxis</i>	0	1	1	0	1	0
<i>Staphylococcus</i>	0	0	1	1	0	0
<i>Streptococcus</i>	1	0	1	1	0	0
<i>Streptomyces</i>	0	0	1	0	1	0
<i>Streptomyces</i>	0	1	1	0	1	0
<i>Symbiobacterium</i>	0	1	1	1	0	0
<i>Synechococcus</i>	0	0	1	1	0	0
<i>Synechococcus</i>	0	1	1	0	1	0
<i>Synechocystis</i>	0	0	1	1	0	0
<i>Syntrophus</i>	0	1	1	0	1	1
<i>Thermoanaerobacter</i>	0	1	1	0	1	0
<i>Thermobifida</i>	0	0	1	0	1	0
<i>Thermosynechococcus</i>	0	1	1	1	0	0
<i>Thermotoga</i>	0	0	1	0	1	0
<i>Thermus</i>	0	0	1	1	1	1
<i>Thiobacillus</i>	0	1	1	0	1	1
<i>Thiomicrospira</i>	0	1	1	0	1	1
<i>Treponema</i>	0	0	1	1	0	0
<i>Treponema</i>	1	0	1	1	0	0
<i>Trichodesmium</i>	0	0	1	1	0	0
<i>Tropheryma</i>	0	0	1	0	1	0
<i>Ureaplasma</i>	1	0	1	1	0	0
<i>Vibrio</i>	0	1	0	1	0	1
<i>Wigglesworthia</i>	0	1	0	1	0	1
<i>Wolbachia</i>	0	0	1	0	1	0
<i>Wolinella</i>	0	0	1	0	1	0
<i>Xanthomonas</i>	0	1	0	1	0	1
<i>Xylella</i>	0	0	0	1	0	1
<i>Yersinia</i>	1	1	0	1	0	1
<i>Zymomona</i>	0	0	1	0	1	0
Total	20	79	121	72	79	64

A utilização da via alternativa de acilação de tRNA para a produção de asparagina livre implica que para haver sua formação é necessário haver primeiramente a ligação de aspartato no *tRNA^{Asn}* para a sua posterior conversão em asparagina. Decidimos testar essa hipótese em *G. diazotrophicus*. Para isso o extrato celular livre de tRNA foi incubado com ATP e asparato marcado radioativamente por 45 minutos. Depois desse tempo uma

cromatografia de camada fina foi realizada para a separação dos aminoácidos. Depois de seca, a placa foi exposta a uma tela amplificadora e a quantidade de asparagina marcada radiotivamente foi determinada. Em células que possuem a via de síntese de asparagina através da asparagina sintetase é possível detectar asparagina formada pois nessa via a formação de asparagina é independente da ligação do aspartato ao tRNA. Já naquelas que tiverem a síntese de asparagina dependente da via alternativa de aminoacilação de tRNA ela não será detectável. Na Figura 4.11 pode-se ver que as células de *E. coli* (que possuem a asparagina sintetase) formaram em média 35 pmoles de asparagina marcada radiotivamente, no entanto no extrato celular de *G. diazotrophicus* não foi possível detectar asparagina radioativa.

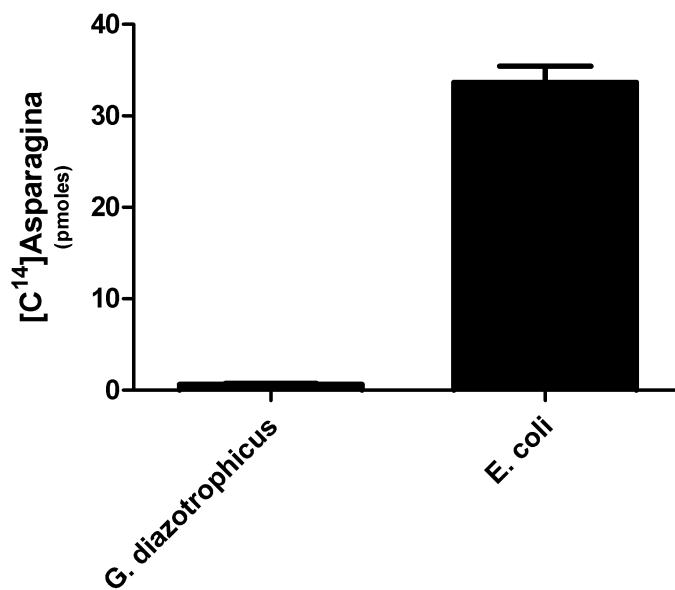


Figura 4.11: Atividade de asparagina sintetase dependente de tRNA. Quantidade de [¹⁴C]asparagina formada pela conversão de [¹⁴C]aspartato por células de *G. diazotrophicus* e *E. coli* (controle positivo) livres de tRNA. Média +/- barra de erros de três experimentos independentes são mostrados.

Para ser possível utilizar a via de transamidação dependente de tRNA para a formação de asparagina, a asparagina ligada ao tRNA (Asn-tRNA) precisa ser hidrolisada. Embora a estabilidade química da ligação Asn-tRNA seja uma das menores de todas as ligações aminoacil-tRNA (Wolfson, 1998), a célula precisa de uma hidrolase para liberar

a asparagina do tRNA para tornar esta via de síntese de asparagina eficiente. Calendar e colaboradores já descreveram e caracterizaram uma destas hidrolases em *E. coli* [Calendar e Berg 1967], o que indica que enzimas com esta função podem estar presentes em *G. diazotrophicus*.

4.3.5 Identificação da via alternativa em fixadores de nitrogênio

Tem sido sugerido [Ibba e Söll, 2000] que a via indireta de formação de asparagina pode ser vantajosa em certos organismos, especialmente naqueles em que o controle dos níveis de asparagina e glutamina são essenciais [Kaneko et al., 2003]. Nós especulamos que a utilização da via alternativa de formação de asparagina poderia ser uma vantagem para a *G. diazotrophicus*, principalmente porque sabemos que a asparagina é um dos principais inibidores do processo de FBN (Tabela 4.5). Foi realizada então uma análise da presença desta via em organismos capazes de realizar o processo de FBN. Foi verificado que embora a capacidade de fixar nitrogênio esteja amplamente distribuída por organismos de classes e gêneros bem distintos entre si, nenhum deles possui uma asparagina sintetase do tipo A (*asnA*), e uma parte considerável (31%) dos fixadores analisados também não possui a asparagina sintetase (*asnB*). Conforme o esperado, todos estes possuíam os genes *gatCAB*. Nestes, a síntese de asparagina seria completamente dependente da *gatCAB*. Interessantemente, todos os que não possuem nenhuma forma de asparagina sintetase retém a capacidade de fixar nitrogênio quando cultivados na sua forma livre, sem associação com planta ou alga.

Tabela 4.5: Identificação da via de síntese de asparagina e da via alternativa de acilação de tRNA em organismos com capacidade de fixar nitrogênio.

Gênero	<i>asnA</i>	<i>asnB</i>	<i>gatCAB</i>	<i>asnS</i>	<i>NDaspS</i>	<i>glnS</i>
de vida livre						
<i>Nostoc</i>	0	0	1	1	0	0
<i>Beijerinckia</i>	0	0	1	0	1	0
<i>Azotobacter</i>	0	1	1	0	1	1
<i>Rhodospirillum</i>	0	1	1	0	1	0
simbióticos						
<i>Anabaena</i>	0	0	1	1	0	0
<i>Frankia</i>	0	1	1	0	1	0
<i>Rhizobium</i>	0	1	1	0	1	0
<i>Sinorhizobium</i>	0	1	1	0	1	0
<i>Mesorhizobium</i>	0	1	1	0	1	0
<i>Bradyrhizobium</i>	0	1	1	0	1	1
endofíticos						
<i>Gluconoacetobacter</i>	0	0	1	0	1	0
<i>Azospirillum</i>	0	1	1	0	1	0
<i>Azoarcus</i>	0	1	1	0	1	1
<i>Herbaspirillum</i>	0	1	1	0	1	0
Total 16	0	10	14	2	12	3

4.3.6 Screening dos mutantes auxotróficos para asparagina

Convencidos da relevância do metabolismo de asparagina no processo de FBN, decidimos investigar as consequências de mutações em genes envolvidos no metabolismo deste aminoácido em no processo de FBN por *G. diazotrophicus*. Para isso foi construída uma biblioteca de mutantes que possuíam uma inserção aleatória do transposon Tn5 no DNA genômico. Foram isolados 5000 mutantes.

Os 5000 mutantes aleatórios foram testados quanto à capacidade de utilização de asparagina com o objetivo de avaliar o efeito da mutação no metabolismo deste aminoácido. Para isso, eles foram crescidos em meio LGI-P suplementado com 10mM de sulfato de amônio e LGI-P suplementado com 10 mM de asparagina. Dos 5000 mutantes analisados, sete mutantes eram auxotróficos para asparagina, ou seja, cresceram no meio suplementado com asparagina, mas não cresceram no meio suplementado com sulfato de amônio (Figura 4.12).

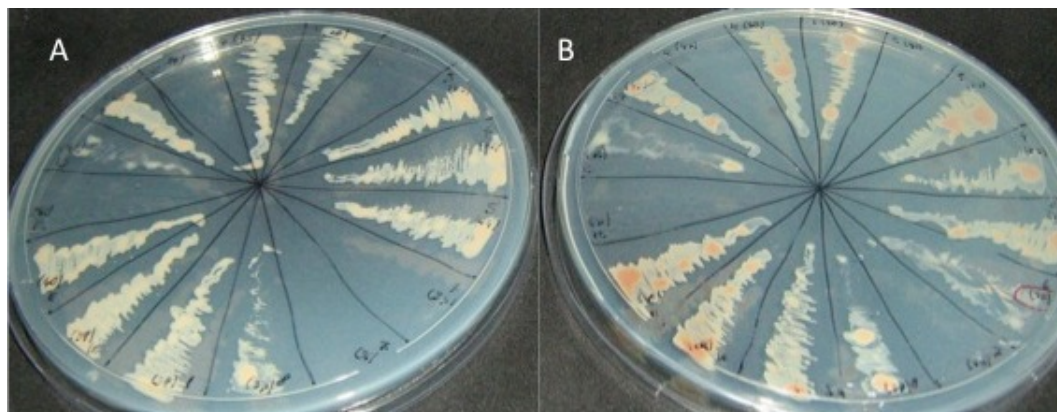


Figura 4.12: Exemplo de placas utilizadas para a seleção dos mutantes auxotróficos para asparagina. Os mutantes que não cresceram na placa “A” (meio LGI-P suplementado com sulfato de amônio), mas cresceram na placa “B” (meio LGI-P suplementado com asparagina) foram selecionadas para experimentos subsequentes.

4.3.7 Identificação e análise dos mutantes

Para identificar o local de inserção do transposon Tn5 foi realizado um PCR invertido (iPCR) a partir do DNA genômico dos mutantes. Os produtos de iPCR foram visualizados em gel de agarose 1% e aqueles que apresentavam uma única banda no gel foram sequenciados. A existência de apenas uma banda no gel do iPCR é um indicativo de que o transposon Tn5 só se inseriu em um local do DNA cromossômico. Os produtos de iPCR sequenciados dos 7 mutantes foram comparados através da ferramenta BLASTX com a sequência do genoma de *G. diazotrophicus*. Foram identificadas inserções em genes com funções relacionadas ao metabolismo de carbono (Fosfoenolpiruvato carboxilase), metabolismo de aminoácidos (Glutamato sintase e Amidofosforibosil transferase), e ao transporte de aminoácidos através da membrana (Permease transportadora de D-metionina). Interessantemente, dos 7 mutantes, 4 eram deficientes na produção da enzima Glutamato Sintase (GOGAT). Os genes estão identificados na Tabela 4.6

Tabela 4.6: Seqüências obtidas dos produtos de iPCR dos mutantes aleatórios.

Gene	Produto	Locus tag
Mutante 4	Fosfoenolpiruvato carboxilase (PEPCase)	GDI3062
Mutante 14	Amidofosforibosil transferase (<i>purF</i>)	GDI3505
Mutante 98	Permease transportadora de D-metionina (<i>metI</i>)	GDI2656
Mutantes 8, 12 e 24	Glutamato sintase, cadeia longa (GOGAT)	GDI2111
Mutante 35	Glutamato sintase, cadeia curta (GOGAT)	GDI2112

Após a identificação dos mutantes fomos finalmente avaliar o efeito de cada mutação no processo de FBN. Para isso as células selvagens e mutantes foram crescidas em meio LGI-P semi-sólido suplementado com 1 mM de sulfato de amônio ou com diferentes concentrações de asparagina (1, 5 e 20 mM). Normalmente, células selvagens crescidas com concentrações maiores que 5 mM de asparagina já teriam a atividade de nitrogenase inibida, mas como os mutantes só crescem em meio com asparagina, nós resolvemos avaliar a atividade de nitrogenase assim mesmo. Nos subítens abaixo pode-se ver uma breve descrição de cada mutante seguido do resultado do teste de redução do acetileno para cada um deles.

Mutante 4 – Fosfoenolpiruvato carboxilase - PEPCase (GDI 3062)

O primeiro mutante analisado possui uma mutação na enzima **Fosfoenolpiruvato carboxilase**. Esta é uma enzima que catalisa a reação anaplerótica mais importante para fornecer intermediários para o ciclo do ácido tricarboxílico que é a formação de oxaloacetato a partir de fosfoenolpiruvato (PEP) e um íon hidrocarbonato [Chollet et al, 1996]. Esta amplamente distribuídas em plantas, algas e bactérias, mas não é encontrada em animais, fungos ou leveduras [Lepiniec et al., 1994, Toh et al., 1994, Gehrig et al., 1998].

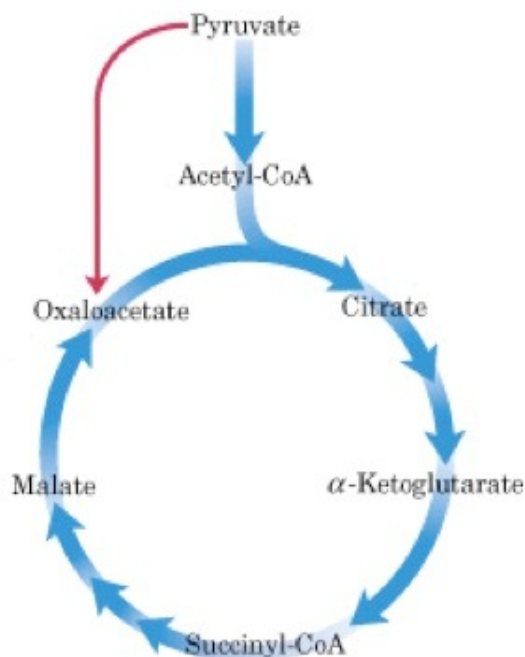


Figura 4.13: Representação gráfica da via metabólica do ciclo do ácido tricarboxílico, onde está destacado (em vermelho) a enzima PEPCase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes.

Sua importância no processo de fixação de nitrogênio em nódulos radiculares de leguminosas também já foi descrita: ela funciona em conjunto com a glutamina, glutamato e asparagina sintetases e com a aspartato amidotransferase, para sintetizar aspartato e asparagina [Gordon et al., 2001]. É relatado que seu principal papel no metabolismo do nódulo é de fornecer carbono para a assimilação do nitrogênio fixado [Schulze et al., 1998].

A fim de se avaliar o efeito de aminoácidos na atividade de nitrogenase na estirpe mutante (MUT4) em comparação com a selvagem PAL5, foi utilizada a técnica de redução de acetileno em meio semi-sólido LGI-P, suplementado com 1mM sulfato de amônio e com diferentes concentrações do aminoácido asparagina (1, 5 e 20mM).

A Figura 4.14 mostra a atividade da nitrogenase na estirpe selvagem PAL5 quando cultivada em meio LGI-P com 1mM sulfato de amônio em contraste com a estirpe mutante. Como era esperado, o mutante 4 foi incapaz de crescer nesta mesma condição. Já em meio semisólido LGI-P contendo 1mM de asparagina como fonte de nitrogênio houve

uma queda de 54% de atividade de nitrogenase na estirpe selvagem, enquanto a estirpe mutante apresentou pouco crescimento, e atividade da nitrogenase muito baixa (Figura 4.14). Na presença de 5 mM de asparagina, PAL5 apresentou atividade da nitrogenase ainda mais baixa (26 nmoles de etileno . mg^{-1} . h^{-1}), enquanto o mutante apresentou crescimento, porém não apresentou atividade de nitrogenase. No meio LGI-P suplementado com 20mM de asparagina tanto a estirpe selvagem quanto o mutante não apresentaram atividade de nitrogenase. O mutante foi incapaz de fixar nitrogênio em qualquer uma das condições avaliadas, provavelmente devido a ausência de uma quantidade de cadeias carbônicas necessárias para a incorporação do nitrogênio fixado.

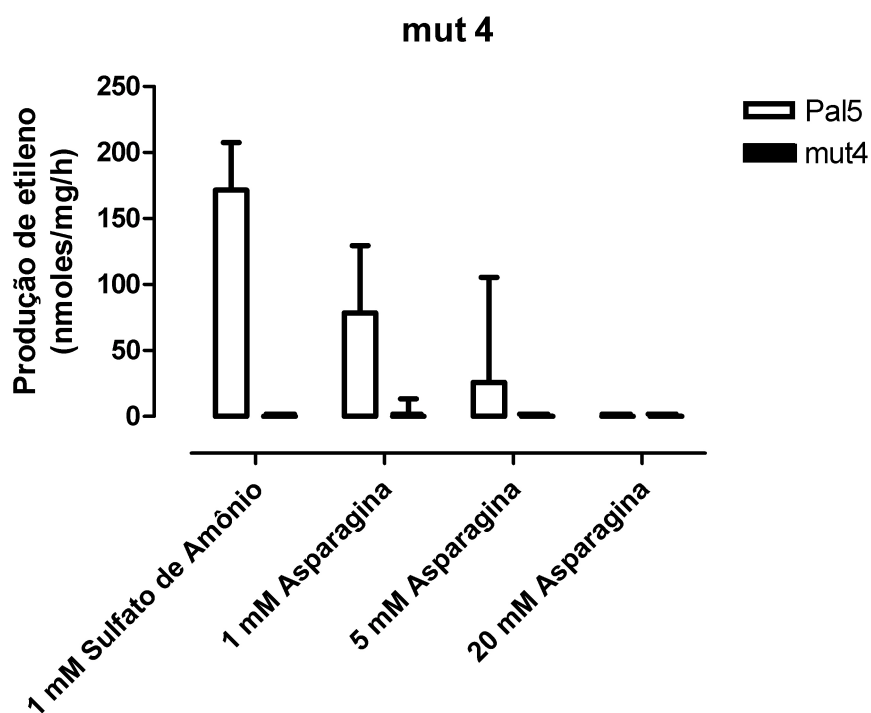


Figura 4.14: Atividade de nitrogenase (em nmoles de etileno . mg^{-1} . h^{-1}) da estirpe selvagem PAL5 e mutante fosfoenolpiruvato carboxilase (MUT4) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados.

De fato já foi demonstrada a influência de mutações no gene da Fosfoenolpiruvato

carboxilase. Em 1998 Schulze e colaboradores criaram mutantes para a enzima em alfafa. Os nódulos destas plantas apresentaram taxas de fixação de CO_2 e N_2 menores que a estirpe selvagem. Estudos posteriores comprovaram uma forte interdependência entre a PEPCase nodular e a atividade de nitrogenase [Schulze et al., 1998].

Mutante 14 – Amidofosforibosil transferase - purF (GDI3505)

O segundo mutante analisado possui uma mutação na enzima **Amidofosforibosil transferase** (ou amidotransferase fosforibosil pirofosfato) que é uma enzima que converte fosforibosilpirofosfato (PRPP) em 5-phosphoribosilamina. A enzima usa o grupo de amônia da cadeia lateral de glutamina. Este é o principal passo na síntese de novo de purinas.

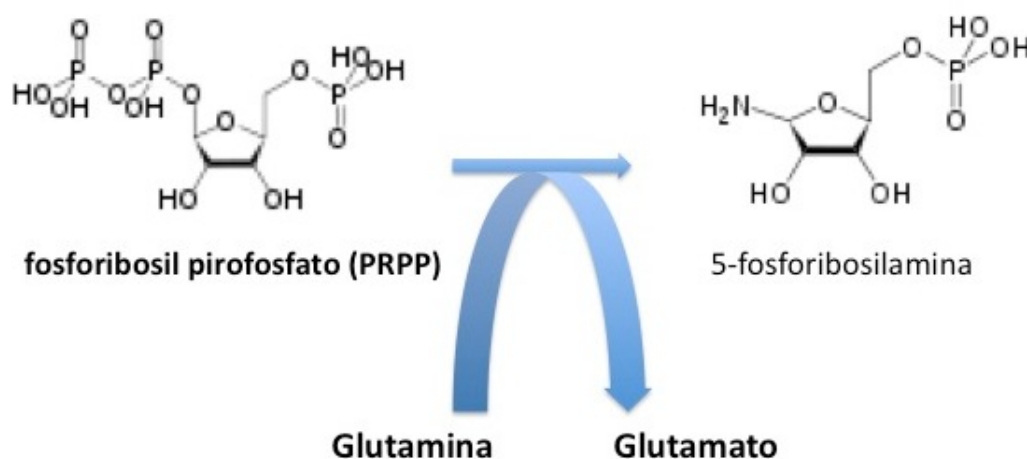


Figura 4.15: Representação gráfica via do metabolismo de glutamato, onde destaca-se (em vermelho) a enzima Amidofosforibosil transferase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes.

O mutante foi avaliado no que diz respeito a sua capacidade de fixar N_2 seguindo a mesma metodologia do experimento 4.3.7. Interessantemente o mutante apresentou uma atividade da nitrogenase maior do que células da estirpe selvagem na presença de 5 ou 20 mM de asparagina como única fonte de nitrogênio. Em células selvagens 20 mM de asparagina normalmente já é suficiente para inibir completamente a FBN, e o fato de que a estirpe mutante não é inibida por essa concentração já é interessante e abre perspectivas

para seu potencial biotecnológico. As taxas de crescimento celular foram alteradas e o mutante tinha a sua taxa de crescimento duas vezes menor, quando comparada à estirpe selvagem.

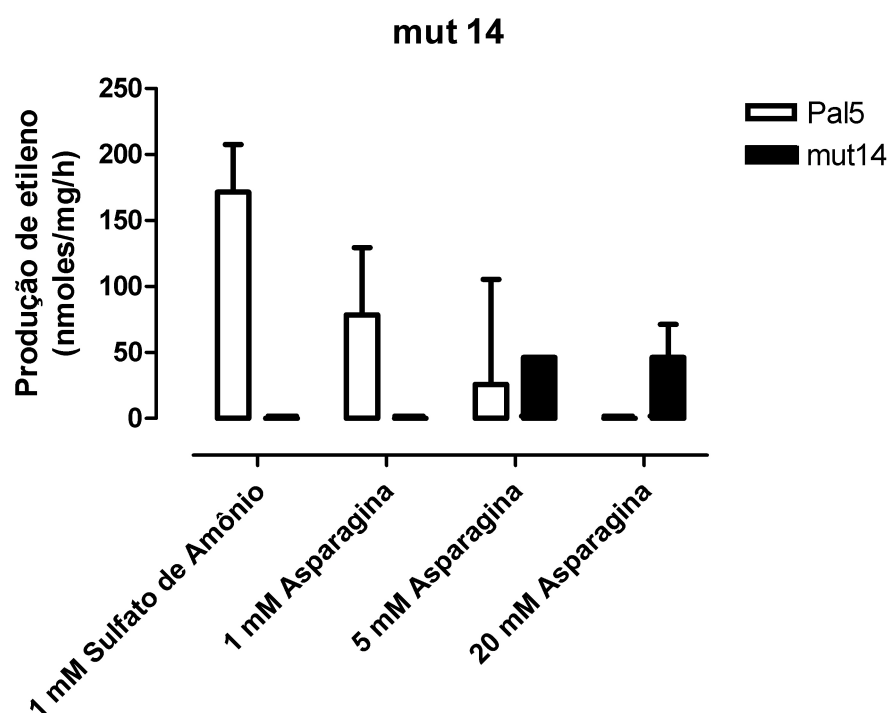


Figura 4.16: Atividade de nitrogenase (em nmoles de etileno . mg⁻¹ . h⁻¹) da estirpe selvagem PAL5 e mutante Amidofosforibosil transferase (MUT14) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média +/- barra de erros de dois experimentos independentes são mostrados.

Mutante 98 – GDI 2656 – Permease do sistema de transporte de D-metionina metI -

Esta subunidade transmembrana (TM) é encontrada em transportadores do tipo ABC (do inglês, ATP-Binding Cassette). Estes tipos de transportadores estão principalmente envolvidos na importação de solutos do meio ambiente. A maioria é específica para um único substrato ou um grupo de substratos relacionados, porém alguns transportadores são mais promíscuos, transportando substratos estruturalmente diversos tais como a his-

tidina/lisina e arginina em *Enterobacteriaceae*.

Foi verificada atividade da nitrogenase na estirpe selvagem PAL5 quando cultivada em meio LGI-P com 1mM sulfato de amônio em contraste com a estirpe mutante (MUT98). Em meio semisólido LGI-P contendo 1mM de asparagina como fonte de nitrogênio houve um aumento de atividade de nitrogenase na estirpe mutante que já apresentou uma atividade maior do que a estirpe selvagem) (Figura 4.17). Na presença de 5 mM de asparagina como única fonte de nitrogênio a atividade de nitrogenase do mutante já era 6 vezes maior do que a da estirpe selvagem e com 20mM de asparagina, enquanto a estirpe selvagem não apresentou atividade de nitrogenase, a atividade de nitrogenase do mutante foi cerca de 200 nmoles de etileno/mg/h).

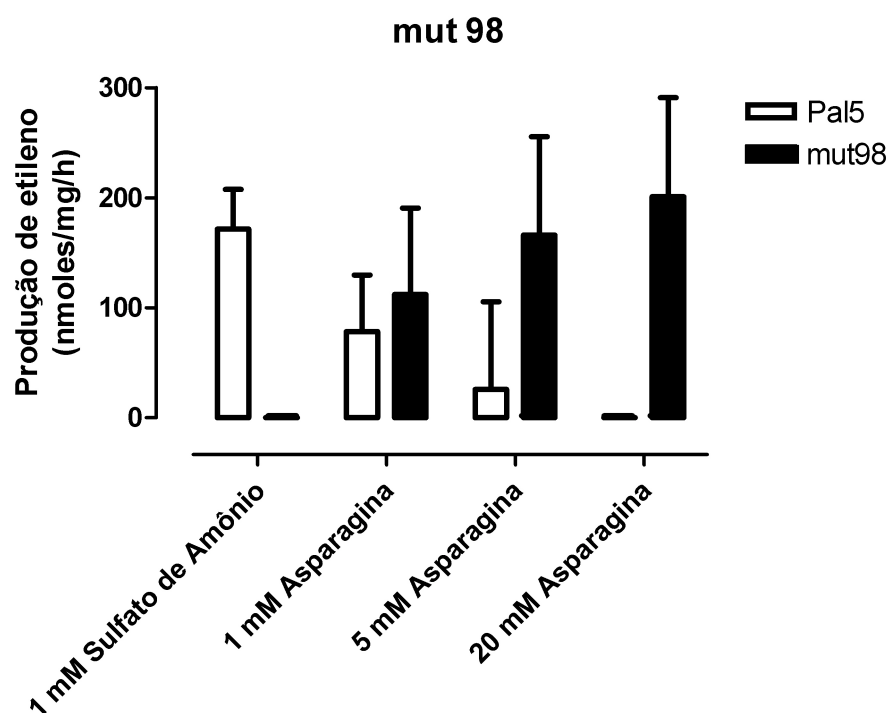


Figura 4.17: Atividade de nitrogenase (em nmoles de etileno \cdot mg^{-1} \cdot h^{-1}) da estirpe selvagem PAL5 e mutante Permease do sistema de transporte de D-metionina metI (MUT98) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média \pm barra de erros de dois experimentos independentes são mostrados.

Mutante 8, 12, 24 – Glutamato sintase, cadeia longa (GDI 2111) e MUT. 35 – GDI 2112 - Glutamato sintase, cadeia curta (GDI 2112)

É a enzima que sintetiza L-glutamato a partir de 2-oxoglutarato e L-glutamina, um passo muito importante na assimilação de amônia. O domínio N-terminal glutaminase catalisa a hidrólise da glutamina em glutamato e amônia, e tem uma dobra semelhante ao de outras glutamina amidotransferases. Em *G. diazotrophicus* a via GS/GOGAT, codificada pelos genes *glnA* e *gltDB*, é a principal via de assimilação de amônia [Alvarez e Drets, 1995]. O gene da glutamato desidrogenase não foi encontrado no genoma de *G. diazotrophicus*, diferentemente do demonstrado em PAL3 [Alvarez e Drets, 1995]. Os fixadores de nitrogênio que não possuem glutamato desidrogenase devem assimilar amônia pela via GS/GOGAT [Kanamori et al., 1988].



Figura 4.18: Representação gráfica da via de metabolismo de nitrogênio, onde destaca-se (em vermelho) a enzima Glutamato sintase. Fonte: Kegg - Kyoto Encyclopedia of Genes and Genomes.

Assim como o mutante para a enzima PEP carboxilase todos os mutantes para a enzima glutamato sintase não apresentaram atividade de nitrogenase em nenhuma das condições testadas. De fato esse fenômeno já era esperado neste tipo de mutante, já que a ausência da glutamato sintase leva a um acúmulo de glutamina na entrada da via de assimilação de amônia, que já seria suficiente para a inibição da atividade de nitrogenase.

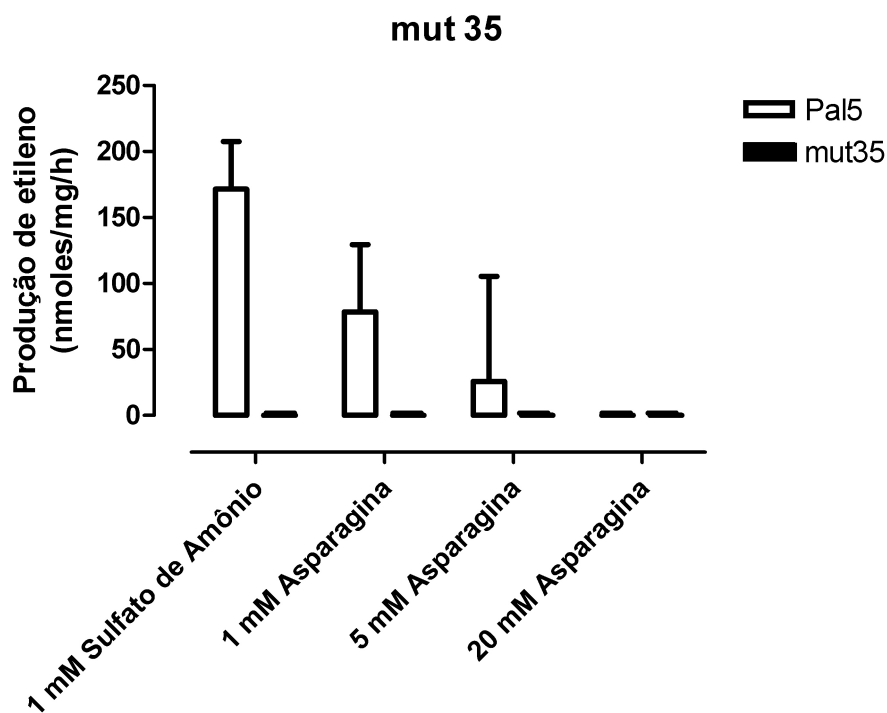


Figura 4.19: Atividade de nitrogenase (em nmoles de etileno \cdot $\text{mg}^{-1} \cdot \text{h}^{-1}$) da estirpe selvagem PAL5 e mutante Glutamato sintase (MUT35) em meio LGIP semi-sólido suplementado com diferentes 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). Média \pm barra de erros de dois experimentos independentes são mostrados.

4.4 Conclusões

A síntese das proteínas é a via através da qual a informação genética é utilizada para a produção de proteínas que expressarão o fenótipo celular. Por isso a acurácia deste processo é um ponto essencial e dependente principalmente do papel exercido pelos três tipos de moléculas de RNA existentes: o RNA mensageiro (mRNA), o RNA transportador (tRNA) e o RNA ribossomal (rRNA). Nesse sentido a acilação do aminoácido com o seu tRNA cognato é um passo crucial para a precisão do processo. Embora seja uma etapa tão importante, pouca ou nenhuma atenção tem sido dedicada ao esclarecimento deste processo em *G. diazotrophicus*. Nosso grupo propôs-se então a estudá-lo, focando principalmente na existência de uma provável via alternativa de acilação de tRNA, seu envolvimento com o processo de produção de asparagina e a influência do metabolismo de asparagina na FBN. Os resultados obtidos até agora nós permitem chegar às seguintes conclusões:

- Uma via alternativa de aminoacilação de tRNA está presente e é ativa em *G. diazotrophicus*.
- Entre as fontes de nitrogênio analisadas, a asparagina, tanto extracelular quanto intracelular, é a principal inibidora da expressão do gene da nitrogenase em *G. diazotrophicus*.
- Uma via alternativa para a formação de asparagina livre intracelular está presente em *G. diazotrophicus* e ela é dependente da via alternativa de aminoacilação de tRNA.
- A via alternativa de formação de asparagina está vastamente distribuída entre os fixadores de nitrogênio de vida livre.
- Foram contruídos 7 mutantes com alterações no metabolismo de asparagina. Todas as alterações foram capazes de influenciar as taxas de FBN em *G. diazotrophicus*.

- Um dos mutantes apresenta um grande potencial biotecnológico pois é capaz de fixar N_2 mesmo quando crescido em altas concentrações de asparagina, como é o ambiente no interior da cana-de-açúcar e merece ser investigado com mais detalhes.

Este é o primeiro estudo onde aborda-se algum aspecto do processo da síntese proteica em *G. diazotrophicus*. Nós comprovamos a existência de vias alternativas para acilação de tRNA e formação de asparagina e demonstramos a influência do metabolismo de asparagina no processo de FBN. Aqui propomos que a existência de uma via alternativa para a produção de asparagina pode ser um mecanismo adotado por fixadores de vida livre para manter o processo de FBN sob maior controle.

Capítulo 5

Materiais e Métodos

5.1 Técnicas de bioinformática

5.1.1 Identificação de ORFs relacionadas ao sistema de transportadores no genoma da *G. diazotrophicus* PAL5

A maioria dos projetos de sequenciamento de genoma completo é geralmente dividido em três fases: o sequenciamento de “shotgun” (WGS), o sequenciamento de finalização e a anotação. Conforme os “reads” são incluídos na montagem do genoma, os “contigs” mudam de uma forma muito dinâmica. Apesar da estratégia do WGS ser um processo de sequenciamento aleatório, a maior parte do genoma é sequenciado, sobrando apenas algumas regiões problemáticas, difíceis de serem clonadas ou sequenciadas, que são chamadas de “gaps”. Desse modo, muito frequentemente, a anotação do genoma é realizada apenas quando terminadas as duas primeiras etapas.

O processo de anotação do genoma do *G. diazotrophicus* PAL5, foi executado no sistema Artemis [Carver et al., 2008], pelos membros integrantes do Projeto Riogene. Foi realizado um processo de anotação manual, onde o primeiro passo consistiu na validação das ORFs, previstas pelo programa GLIMMER [Delcher et al., 1999].

Nestas situações, a seleção das ORFs corretas foi baseada nos resultados de Blast [Altschul et al., 1998], onde dentre as ORFs sobrepostas, as que apresentaram melhores

alinhamentos em relação aos bancos de dados foram validadas. Dentre as variantes de Blast [Wu-blast, 2006] utilizadas no processo de anotação, o Blastp foi utilizado como ferramenta inicial, pois compara sequências de aminoácidos, contra sequências de proteínas depositadas nos bancos de dados públicos, realizando uma análise menos estridente que outras ferramentas, tais como o Blastn (utilizado para comparação de sequências de nucleotídeos), devido ao fato de que os mais variados organismos podem utilizar códons diferenciados, para codificar o mesmo aminoácido.

A inferência de prováveis funções das proteínas codificadas pelo genoma de *G. diazotrophicus* PAL5, com base nos resultados de Blastp obedeceu os seguintes critérios:

- o alinhamento deve ter e-value inferior a 10^{-5} ;
- porcentagens de identidade superiores a 30%;
- alinhamentos de 60% ou mais da sequência da ORF foram considerados suficientes para caracterizar homologia com sequências de proteínas depositadas em bancos de dados. Quando não houve detecção de sobreposição de sequências de ORFs, e mesmo assim o resultado de Blast indicou uma similaridade muito baixa ou nenhuma em relação aos bancos de dados, estas sequências foram classificadas como sequências de proteínas hipotéticas.

Posteriormente, realizou-se uma análise na base de dados PFAM [Bateman et al., 2004], que apresenta uma coleção de famílias e domínios protéicos. O PFAM possui sequências conservadas de cerca de 6200 famílias protéicas, das quais aproximadamente 75% das proteínas encontradas no SWISS-PROT [Boeckmann et al., 2003], possuem pelo menos um domínio do PFAM, e cada resultado relatado no PFAM acompanha um comentário sobre o domínio encontrado, facilitando a identificação da possível função da ORF. Após a predição da possível função das ORFs, estas foram validadas e as suas respectivas categorias e subcategorias funcionais identificadas, assim como o nome do gene, número de EC (Enzyme Category), número de COG (Clusters of Orthologous Groups), organismos relacionados, e comentários adicionais dos pesquisadores acerca da ORF.

Adicionalmente, na tentativa de determinar prováveis funções para ORFs de algumas proteínas hipotéticas, foram analisados os alinhamentos de sequências do *G. diazotrophicus* PAL5 (Blastp), que apresentaram alinhamentos com e-values inferiores a 10^{-3} , e alinhavam com domínios protéicos conservados depositados no PFAM. As proteínas hipotéticas com valores de e-value superiores a 10^{-3} nos seus alinhamentos continuaram a receber esta classificação.

A maioria, se não todas, as proteínas que estão localizadas na membrana externa das bactérias, por meio de uma via secretora, contem um peptídeo de sinal. O programa SignalP [Emanuelsson et al., 2007] analisa a sequência de aminoácido de uma ORF para peptídeos sinais, e os primeiros 60 aminoácidos N-terminais de cada ORF foram analisados usando a base de dados de bactérias Gram negativas.

Outro programa utilizado foi o Psort, que determina a localização de uma proteína com base na sequência de peptídeo sinal, no reconhecimento de segmentos transmembranares e na análise de sua composição de aminoácidos. Considera-se que um produto de uma ORF é uma proteína de membrana se os dados do resultado prevêm com um valor de segurança de 0,5 ou maior.

Todas as possíveis proteínas de membrana foram adicionalmente, analisadas por alinhamentos comparativos com bancos de dados para proteínas de transportes, disponibilizados pelos grupos: i) Dr. Milton Saier Jr. (University of California, San Diego, UCSD/USA) para utilizar o banco de dados do TCDB [Saier et al., 2006](Transporter Classification System – também disponível em: <http://www.tcdb.org/>) ; ii) Drs. Ian T. Paulsen e Ren Quen do Insititute J. Craig Venter e The Institute for Genomic Research – TIGR para anotação automática e comparativa do sistema de transporte de *G. diazotrophicus* com outros genomas sequenciados e disponíveis no TransportDB – Genomic Comparisons of Membrane Transport System: <http://www.membranetransport.org/>; e, iii) Dr. Yves Quentin da Université Paul SABATIER-CNRS/France para anotação automática no ABCdb (Archaeal and Bacterial ABC transporter database; disponível em : <http://www-abcdb.biotoul.fr/>; [Fichant et al., 2009]).

5.1.2 Análise dos clusters

A clusterização foi feita utilizando-se os valores de mRNA quantificados por qPCR de células crescidas em condições de fixação ou não fixação de nitrogênio. A expressão dos genes *sodA*, *katE*, *kat*, *katC*, *gorA*, *gorB* e *nifD* nestas condições foi determinada nos tempos de 18, 24, 36, 48 e 72 horas de cultivo. Os genes foram agrupados então de acordo com seus padrões de expressão ao longo do tempo e condições de cultivo. Desta maneira genes que respondem da mesma maneira a diferentes condições são agrupados. A análise dos clusters foi realizada utilizando-se o pacote MINITAB v.14 package, usando-se “Euclidian distances” e os métodos “Ward linkage”, “Single linkage”, “Complete linkage”, “Average linkage”, “Centroid linkage” ou “Median linkage”.

5.2 Técnicas de biologia molecular

5.2.1 Extração de DNA genômico

A extração do DNA total foi realizada de acordo com o método CTAB (Brometo de cetiltrimetilamônio) conforme descrito por Doyle e Doyle [Doyle e Doyle, 1987]. Para a extração do DNA total as estirpes mutantes e selvagem de *G. diazotrophicus* foram cultivadas em 5 ml de meio DYGS líquido até uma A_{600nm} de aproximadamente 1,0. Dois mL da cultura líquida de *G. diazotrophicus* em fase de crescimento logarítmico em meio DYGS foram precipitados por centrifugação a 18.514 g, 4°C por 10 minutos. As células foram lavadas uma vez com 500 µl de tampão TE (10 mM de Tris pH 8,0 contendo 1 mM de Na_2EDTA), centrifugadas novamente, e ressuspensas em 567 µL de tampão TE. Em seguida, 30 µL de uma solução de SDS (10%), mais 3 µL de proteinase K (20 $mg.ml^{-1}$) foram adicionados. O material foi homogeneizado e incubado a 37 °C durante uma hora. Após a incubação foi adicionado 100 µL de uma solução de NaCl 5 M, seguidos de 80 µL de uma solução de CTAB (brometo de cetiltrimetilamônio 10% em NaCl 0,7 M). Após homogeneização, as amostras foram incubadas por 10 minutos a 65°C, seguidas por

1 h sob agitação leve, a temperatura ambiente. O primeiro passo da extração foi efetuada pela adição de 1 volume de uma solução de fenol:clorofórmio:álcool isoamílico (25:24:1, vol./vol.), homogeneização de 2 minutos por inversão dos tubos e centrifugação por 10 minutos a 18.514g a 4°C. Após a transferência do sobrenadante para um novo tubo, foram adicionados 5 µL de RNase (5 mg.ml⁻¹), e a amostra incubada a 37°C por 1 hora. Após incubação, foram realizados dois passos de purificação: o primeiro passo com um volume de fenol:clorofórmio:álcool isoamílico (25:24:1, vol./vol.) e o segundo com 1 volume de clorofórmio:álcool isoamílico (24:1, vol./vol.). O material foi homogeneizado durante 2 minutos, centrifugado (10 minutos, 18.514 g, 4°C) e transferido para novos tubos, nas duas etapas. O passo seguinte corresponde a adição de 0,6 volumes de isopropanol 100% (5 minutos de incubação), seguido de duas etapas de lavagem com etanol 70%, secagem e solubilização em 40 µL de água.

5.2.2 Eletroforese de DNA

Os produtos das reações de PCR e o DNA cromossossômico foram eletroforéticamente separados em gel de agarose 0.8 ou 1% (w/v) preparado em tampão TAE (40 mM Tris-acetato, 1 mM EDTA (pH 8,0), contendo brometo de etídio na concentração final de 0,5 µg/mL. Após a corrida o DNA foi visualizado sob luz ultravioleta (Sambrook et al., 1989).

5.2.3 Extração do RNA total

A extração de RNA foi realizada utilizando o reagente TRIzol [Chomczynski e Sacchi, 1987] aquecido a 50°C. Alíquotas de 1,5 mL de cultura de células foram centrifugadas, o sobrenadante descartado, e homogeneizadas em 1 mL do reagente, uma solução monofásica de fenol e isotiocianato de guanidina, que garante a integridade do RNA durante o processo de lise celular. A seguir, a adição de clorofórmio e uma rápida centrifugação dividem o material em duas fases: uma fase orgânica e uma fase aquosa, na qual se encontra o RNA. Após a passagem da fase aquosa, contendo o RNA desejado, para um novo tubo, este pode ser recuperado através de uma precipitação com isopropanol. Neste experimento,

após a precipitação, o RNA foi lavado com etanol 70% e ressuspenso em 10 μL de água para a dosagem no espectrofotômetro.

O RNA total purificado foi dosado e sua qualidade monitorada por medida da absorção de luz à 260, 280 e 320 nm. 2,5 μg de RNA foram tratadas com 1 U de DNase I (Invitrogen) de acordo com o protocolo do fabricante. Após 30 minutos, foi acrescentado 1 μL EDTA 25 mM e as amostras foram incubadas a 65 °C, por 10 minutos, para inativação da enzima. 1 μL de cada uma das amostras de RNA tratadas com DNase foram testadas para garantir que não havia contaminação genômica através de PCR convencional.

5.2.4 Eletroforese de RNA

Os produtos das extrações de RNA foram eletroforéticamente separados em gel de agarose 0.8 ou 1% (w/v) preparado em tampão TAE (40 mM Tris-acetato, 1 mM EDTA, pH 8), contendo brometo de etídio na concentração final de 0,5 $\mu\text{g}/\text{ml}$. Após a corrida o RNA foi visualizado sob luz ultravioleta [Sambrook e Russell, 1986].

5.2.5 Síntese do cDNA

A síntese da primeira fita de DNA complementar (cDNA) foi realizada por transcrição reversa do RNA obtido, utilizando o kit "High Capacity cDNA Reverse Transcription Kit (Applied Biosystems)", conforme o protocolo do fabricante. Todo o volume obtido no tratamento com DNase foi utilizado para a reação de síntese de cDNA. A reação ocorre com a adição de dNTPs, "random primers", tampão e a enzima transcriptase reversa do KIT ao RNA já tratado. Após 2 horas a 37 °C, a reação é interrompida através da desnaturação da enzima à 75°C, por 10 minutos. Os cDNA foram estocados à -20 °C até serem utilizados.

5.2.6 Amplificação do cDNA por reação em cadeia da polimerase (PCR)

Para conferir a amplificação do segmento desejado do cDNA, todos os pares de iniciadores foram testados através de PCRs convencionais antes do teste na máquina de PCR em tempo real. As reações continham tampão da enzima Taq, dNTPs 0,2 mM, os oligonucleotídeos 1 μ M, 1 μ l de cDNA, água e 1 U da enzima Taq DNA polimerase numa reação de 25 μ l. Todos os genes foram amplificados numa reação com a seguinte programação: 94°C por 4 min; 40 ciclos de 30 segundos a 94°C, 30 segundos a 60°C, 30 segundos a 72°C; e um último ciclo de 15 min a 72°C num termociclador Gene Amp PCR System 2400 (Applied Biosystems) e analisados em gel de agarose 2%.

5.2.7 PCR em tempo real ou PCR quantitativo (qPCR)

A análise da expressão relativa dos genes de interesse foi realizada utilizando a técnica de qPCR.

As reações de PCR em tempo real foram conduzidas em um equipamento Applied Biosystems 7500 Real Time PCR System (Applied Biosystems). Como sistema de detecção foi utilizado o SYBR Green (Power SYBR® Green PCR Master Mix – Applied Biosystems), um reagente que emite fluorescência ao se ligar inespecificamente a qualquer dupla fita de DNA. Por causa dessa característica não específica do SYBR Green, todas as amostras utilizadas nos ensaios de PCR em tempo real foram previamente tratadas com DNase, com o intuito de eliminar qualquer tipo de contaminação com DNA genômico que atrapalhasse a quantificação. As amostras de RNA tratadas com DNase foram testadas para garantir que não havia contaminação genômica através de PCR convencional. Todos os experimentos de quantificação foram realizados com triplicatas biológicas, e cada uma destas com triplicatas técnicas. Para cada gene estudado foi utilizado um par de oligonucleotídeos específico (Tabela 5.1). Antes de realizar a quantificação, cada oligonucleotídeo foi testado por PCR convencional para estabelecer as condições ótimas da reação. Poste-

riormente, os oligonucleotídeos eram submetidos a uma curva de eficiência. Neste teste, um par de oligonucleotídeos (1 μM final) era utilizado em reações com diferentes diluições de um mesmo cDNA (1, 1:10, 1:100 e 1:1000). A curva gerada neste procedimento indica a eficiência de reação daquele determinado oligonucleotídeo naquela amostra, nos permitindo garantir que a quantificação será válida dentro do método escolhido para o processamento dos resultados (Método do $\Delta\Delta\text{Ct}$).

Como o SYBR Green marca qualquer dupla fita de DNA, a análise do branco de reação (reação que não contém cDNA) nos permite determinar a temperatura de melting (T_M , do inglês melting temperature) do dímero de oligonucleotídeo formado na ausência de amplificação. Essas temperaturas nos permitem avaliar, após cada experimento, se a reação produziu o resultado esperado para aquele determinado par de oligonucleotídeos.

A reação foi feita em placas de 96 poços (MicroAmp Optical 96-Well Reaction Plate, da empresa Applied Biosystems). Em cada poço foram colocadas 6,25 mL da mistura de reação da primeira fita de cDNA (diluída em água milliQ estéril), 12,5 μL da solução mix do kit e 6,25 μL de uma mistura dos dois oligonucleotídeos para a concentração final de 1 μM . O gene 23S, que codifica uma proteína ribossomal, foi utilizado como controle endógeno para a normalização em todos os experimentos.

A Tabela 5.1 apresenta a listagem de todos os oligonucleotídeos utilizados nesse trabalho, mostrando a sequência de nucleotídeos de cada um deles.

Tabela 5.1: Primers qPCR

Nome do gene	Locus tag	sequência
(23S)		F:5'-AAAGCCGGATCAATCCGTTA-3' R:5'-AAGCCGTAGTCGATGGAAAC-3'
Nitrogenase (<i>NifD</i>)	GDI0438	F:5'-TTCGGCCACAACGACGACTATCA-3' R:5'-TGCCCATCTTCTGCACGGGATATT-3'
Superóxido dismutase (<i>sodA</i>)	GDI2168	F:5'-CCCCGCCCATGCAAT- 3' R:5'-AGGGTGGGCAATTTCGAAAG-3'
Catalase tipo E (<i>KatE</i>)	GDI0079	F:5'-GTTCGAAGTCACCCACGATGT -3' R:5'-CCCGGTTGCAGGAATTTG-3'
Catalase (<i>Kat</i>)	GDI0467	F:5'-ACGATATTCCGGTCTTCGTC-3' R:5'-CAGGTCGCGAAATTCCTG-3'
Catalase tipo C (<i>KatC</i>)	GDI2359	F:5'-TTCATCGTCGTGACCTTTTG-3' R:5'-TTCCATTCCGAAAATCG-3'
Glutaciona redutase tipo A (<i>gorA</i>)	GDI2216	F:5'TGTGTCCCAAGAAGCTGATG3' R:5'GATGTTCCAGCCGAAACCAT3'
Glutaciona redutase tipo B (<i>gorB</i>)	GDI2280	F:5'-GAGGCGATGCTGTTTACCC-3' R:5'-TCCCAATAGGACGAAACCTG-3'
Glutamate Sintase (GOGAT)	GDI2114	F:5'GCAACAATATTCCCGACTGGCTGA-3' R:5'TTCTCGATCACGCAATTGCCTTCG-3'
Aspartil-tRNA sintetase (AspRS)	GDI2235	F:5'ACCATGCATCCCTATCGTACCCAT-3' R:5'AATGACGATCTGGGTCATGCCGAA-3'
Glutamil-tRNA sintetase (GluRS)	GDI2214	F:5'ACGAGTATGTCGATGCCCTGCATA-3' R:5'TGCTCGCTCTCGAAACACGGATA-3'
Amidotransferase subunidade A (<i>gatA</i>)	GDI2982	F:5'ATCAAGGACCTGTTCTGCACCGA-3' R:5'AAATTGGTCTTGCCCACGAACACC-3'
Amidotransferase subunidade B (<i>gatB</i>)	GDI2981	F:5'TGGAAATCGTCAGCGAACCCGATA-3' R:5'GCGCATCGAACCTTCTTCCATGTT-3'
Amidotransferase subunidade C (<i>gatC</i>)	GDI2983	F:5'TCTCGGCTGGATCGAGCAATTGAA-3' R:5'AACCACCTTCGGCACGGTATAGAA-3'

As reações foram realizadas nas seguintes condições: 95 °C por 10 minutos, 40 ciclos de duas temperaturas: 95 °C por 15 segundos e 60 °C por 1 minuto, seguido de uma curva de desnaturação.

5.2.8 Cálculo da quantificação relativa

O programa ABI Prism 7500 Sequence Detection software versão 1.2.3 (Applied Biosystems) acompanha a reação de PCR em tempo real, gerando um gráfico correspondente ao logaritmo do valor da fluorescência medida em cada poço pelo número do ciclo corres-

pondente. Uma linha de corte (denominada Threshold) é delimitada em um ponto onde todas as curvas (de todos os poços do experimento) estejam em sua fase exponencial.

O ciclo do PCR no qual a curva de amplificação de um determinado gene cruza a linha de corte é denominado Ct (Ciclo de acordo com o Threshold). De acordo com a quantidade inicial de DNA presente em cada amostra analisada, esta amostra irá levar um número maior ou menor de ciclos para que sua curva de amplificação alcance a linha de corte (threshold). Deste modo, amostras com pouco DNA apresentarão um Ct elevado, enquanto amostras com mais DNA apresentarão um Ct mais baixo.

O método que utilizado para realizar a quantificação relativa dos genes foi o Comparative Ct Method, também chamado de Método do $\Delta\Delta Ct$ [Livak e Schmittgen, 2001, Pfaffl, 2001], utilizando o gene 23S como normalizador endógeno. O detalhamento dos cálculos utilizados para a análise dos dados encontra-se explicado nos manuais da Applied Biosystems (Método do $\Delta\Delta Ct$ - User boletim #2, Applied Biosystems, 2001). A expressão relativa final de cada gene foi estabelecida pela média entre as expressões relativas dos diferentes experimentos.

A validação do Método do $\Delta\Delta Ct$ foi efetuada de acordo com o manual da Applied Biosystems (User boletim #2, Applied Biosystems, 2001), através da comparação entre as curvas de eficiência de cada par de oligonucleotídeo.

Tabela 5.2: Eficiência dos oligonucleotídeos. Nome do gene, temperatura de desnaturação (TM), o R^2 da curva de eficiência e a eficiência de amplificação de cada um dos oligonucleotídeos utilizados nesse trabalho. A eficiência foi calculada utilizando a fórmula $E = 10^{(-1/slope)}$ e $\%E = (E - 1) * 100$

Nome do gene	TM (°C)	R^2 Curva de eficiência	Slope	Eficiência (E)	Eficiência (%)
(<i>23S</i>)	71,8	0,9981	-3,40	1,96	96
Nitrogenase (<i>NifD</i>)	75,2	0,9910	-3,43	1,95	95
Superóxido dismutase (<i>sodA</i>)	77,6	0,9873	-3,45	1,94	94
Catalase tipo E (<i>KatE</i>)	78,7	0,9934	-3,32	2,00	100
Catalase (<i>Kat</i>)	80,1	0,9934	-3,50	1,93	93
Catalase tipo C (<i>KatC</i>)	77	0,9955	-3,54	1,91	91
Glutathiona redutase tipo A (<i>gorA</i>)	72,2	0,9843	-3,20	2,05	105
Glutathiona redutase tipo B (<i>gorB</i>)	81,4	0,9947	-3,47	1,94	94
Aspartil-tRNA sintetase (AspRS)	88,5	0,9821	-3,30	2,00	100
Glutamil-tRNA sintetase (GluRS)	86,7	0,9980	-3,45	1,94	95
Amidotransferase subunidade A (<i>gatA</i>)	87	0,9832	-3,24	2,03	103
Amidotransferase subunidade B (<i>gatB</i>)	85,3	0,9954	-3,50	1,93	93
Amidotransferase subunidade C (<i>gatC</i>)	84,6	0,9945	-2,87	1,91	92

5.2.9 Obtenção dos mutantes aleatórios

Os mutantes aleatórios foram obtidos de acordo com o protocolo utilizando o Kit EZ::TN™ <KAN-2> Tnp Transposome™ (Epicentre) que contém o gene Tn903 (*kanR*) de resistência à canamicina na concentração de $200 \mu\text{g} \cdot \text{mL}^{-1}$ [Rouws et al., 2007]. O EZ::TN Transposome forma um complexo estável com a enzima Transposase e pode ser transferido para células vivas. A Transposase é ativada por Mg^{+2} no interior da célula receptora tendo por resultado a inserção aleatória do transposon no DNA genômico.

5.2.10 Caracterização molecular dos mutantes aleatórios (PCR invertido)

Para identificar o local de inserção do transposon nos mutantes aleatórios foi utilizada a técnica de PCR invertido ou iPCR [Martin e Mohn, 1999] baseado no protocolo de Rouws et al. (2007). O DNA total (1 µg/mL) dos mutantes foi digerido com a enzima PstI (37,5 U/reação) ou PvuI (10U/ reação). Após a digestão, o DNA foi tratado com DNA ligase T4 (100U) e utilizados como molde para as reações de iPCR com os primers FP-1 (5'-ACCTACAACAAAGCTCTCATCAACC-3') e RP-1 (5'-GCAATGTAACA TCAGAGATTTTGAG-3') do Kit EZ::TN™ <KAN-2> Tnp Transposome™ (Epicentre), ou com os primers PVU-LEFT (GAAAAACAGCATTCCAGGTATTAGA), PVU-RIGHT (CAAAAAGTTTATGCATTTCTTTCCAG) [Rouws et al., 2007]. Os produtos de iPCR foram então seqüenciados e comparados com as seqüências de *G. diazotrophicus*.

5.2.11 Purificação de DNA para reação de sequenciamento

A purificação de DNA para reação de sequenciamento foi feita utilizando-se o kit “Wizard® SV Gel and PCR Clean up system” da Promega seguindo as instruções do fabricante.

5.2.12 sequenciamento das regiões flanqueadores do transposon

EZ-Tn5™ <KAN- 2>Tnp transposome™

O produto gerado pelo PCR invertido foi purificado como descrito no item 5.2.11. As reações de sequenciamento foram realizadas pela adição do reagente “DYEnamic ET terminator reagent premix” (GE Healthcare), de oligonucleotídeos iniciadores Kan-2 FP-1 ou Kan-2 RP-1 e água ao produto do iPCR purificado. A reação de amplificação consistiu de 30 ciclos de desnaturação à 95°C por 20 segundos, anelamento a 50°C por 15 segundos e extensão a 60°C por 1 minuto. Em seguida, o produto dessa reação foi precipitado por meio da adição de 1 µl de NH₄Ac 7,5 M e 27,5 µl de etanol 100%, sendo incubado

à 4°C por 16 horas. Logo após, o produto foi recuperado por centrifugação (3.220 g, 30 minutos, 4°C) e o sobrenadante descartado. O DNA precipitado foi lavado com 150 µl de etanol 70% e após a evaporação total do etanol 70%, o produto foi redissolvido em 7,5 µl de tampão de amostra para sequenciamento. Após duas horas a 4°C, as amostras foram injetadas no seqüenciador MegaBACE 1000 (GE Healthcare).

5.3 Técnicas de microbiologia e bioquímica

5.3.1 Estirpes, meios e condições de cultivo

Estirpes

A cepa de *G. diazotrophicus* utilizada em todos os experimentos foi a cepa sequenciada PAL 5 (BR11281) obtida da coleção de culturas da EMBRAPA.

Meios de cultura

As estirpes selvagem e mutantes de *G. diazotrophicus* foram cultivadas em meio DYGS [Rodrigues Neto et al., 1986], Batata-P sólido [Cavalcante e Döbereiner, 1988] ou LGI-P [Döbereiner et al., 1995]. Para ensaios de atividade de nitrogenase as células de *G. diazotrophicus* foram cultivadas em meio LGI-P semi-sólido ou líquido contendo ou não diferentes fontes de nitrogênio conforme indicado em cada experimento. Para se obter os meios sólido e semi-sólido, os meios líquidos foram acrescidos de ágar de acordo com as indicações abaixo. Os meios foram suplementados com 200 µg. mL⁻¹ de canamicina quando necessário.

Tabela 5.3: Quantidades de ágar acrescidos aos meios líquidos a fim de se obter meios semi-sólidos ou sólidos.

	Concentração de ágar (g/L)	
meio	sólido	semi-sólido
DYGS	15	-
LGI-P	25	1,6
Batata-P	25	-

Condições de cultivo

As células de *G. diazotrophicus* estirpe PAL5 (BR11281) foram crescidas em meio LGI-P com agitação de 100 rpm por 48h a 30°C. Depois de 3 “rounds” de cultivo sucessivos, as células foram recuperadas por centrifugação a 12000g por 2 minutos e ressuspensas no meio indicado, de acordo com o experimento.

Basicamente, foi estabelecido que as condições de fixação são aquelas em que a célula é crescida somente com 1 mM de sulfato de amônio no meio de cultura e as condições de não fixação aquelas em que a célula é crescida com 20 mM de sulfato de amônio no meio de cultura.

A resistência de *G. diazotrophicus* ao estresse oxidativo foi investigada pela adição de Paraquat (PQ), um gerador de radicais superóxido, em concentração final de 0,5 ou 5 mM, em meios suplementados com 1 ou 20 mM de sulfato de amônio, após 24h de cultivo.

5.3.2 Determinação da produção de ROS em *G. diazotrophicus*

A presença de ROS foi analisada por microscopia de fluorescência de células crescidas em meio líquido LGI-P sob condições de FBN ou não. As culturas de *G. diazotrophicus* foram concentradas 3X e incubadas por 10 minutos na presença de 5 µM do probe sensível à ROS CM-H2DCFDA (Invitrogen – Molecular probes), lavadas com meio LGI-P novo e imediatamente visualizadas por visualização direta (sem agentes fixadores). As imagens foram obtidas em um microscópio Zeiss Axioskop com uma Axiocam MRC5 usando-se o conjunto de filtros Zeiss-09 (excitação – BP 450-490; beam splitter – FT 510; emissão – LP 515). Aumento de 630X.

5.3.3 Atividade de catalase

O consumo de peróxido de hidrogênio foi monitorado a 240nm por 1 minuto à temperatura ambiente em um espectrofotômetro GBC UV/VIS 920. As reações foram feitas em uma cubeta de 1mL na presença de 40µL de tampão Tris-HCl 100mM pH 8, pellet de 1mL de células frescas ressuspendido em 150uL de tampão Tris-HCl 100mM pH 7,8 sonificado 10 vezes em pulsos de 5 segundos, 900µl de peróxido de hidrogênio 9mM. Os brancos foram feitos somente com tampão e peróxido de hidrogênio. O controle negativo foi feito da mesma maneira que a reação principal, mas o pellet de células foi fervido 2 vezes por 15 min antes de ser adicionado à reação. Uma unidade de catalase é definida como a quantidade de enzima necessária para decompor 1µmol de H_2O_2 x min^{-1} .

5.3.4 Preparação de extrato celular livre de tRNA

10 g de células de *G. diazotrophicus* ou *E. coli* (usada como controle positivo) foram centrifugadas na fase logarítmica de crescimento e ressuspendidas em 20 mL do Tampão A (50 mM Hepes-KOH, pH 7,2/ 10 mM $MgCl_2$ / 10% glicerol/ 5 mM DTT) com 0,5 mM de fenilmetilsulfonil fluorido (PMSF, do inglês phenylmethylsulfonyl fluoride). A suspensão foi sonicada e depois de uma centrifugação de baixa velocidade (8700 x g) para remover os restos de células o sobrenadante foi centrifugado por 90 min a 100000 x g. O sobrenadante foi aplicado então em uma coluna Q-Sepharose (Pharmacia) equilibrada com tampão A contendo 300 mM de NaCl.

5.3.5 Atividade de asparagina sintetase

Os extratos celulares de *G. diazotrophicus* ou *E. coli* DH5 livres de tRNA foram preparados como descrito no item 5.3.4. Os meios reacionais continham 50 mM Hepes-KOH (pH 7,2), 25mM de KCl, 10mM de $MgCl_2$, 4 mM de DTT, 4 mM de ATP, 2,5 mM de glutamina, 3,5 mM [^{14}C] aspartato e o extrato protéico livre de tRNA de *G. diazotrophicus* ou *E. coli*. A reação foi feita a 30°C para *G. diazotrophicus* e 37°C para *E. coli*. Depois de

45 minutos de incubação alíquotas de foram separadas por cromatografia de camada fina (TLC, do inglês thin-layer chromatography) em cululose (Sigma) com o solvente chloroformymethanolywateryammonium hydroxide (6:6:1:2). Depois de seca, a placa foi exposta a uma placa phosphorimaging (Fuji) para detectar os aminoácidos marcados. A localização dos aminoácidos foi realizada correndo-se 20 nmol do aminoácido correspondente não marcado seguido de visualização utilizando-se uma solução de ninhidrina 0,2% em etanol.

5.3.6 Atividade de nitrogenase

A Atividade de Redução de Acetileno (ARA) foi determinada em meio de cultura semi-sólido LGI-P, utilizando-se 3 repetições. Uma colônia isolada foi inoculada no meio líquido DYGS e cultivada sob agitação a 200 rpm por 48 h a 30°C. As células foram lavadas com solução salina e a A_{600} ajustada para 0,5. Uma alíquota de 20 μ L dessa suspensão foi transferida para frascos com capacidade de 10 mL contendo 5 mL de meio de cultura semi-sólido contendo 1 mM de sulfato de amônio ou diferentes concentrações de asparagina (1, 5 ou 20 mM). As culturas foram incubadas a 30°C por 72h. Neste momento os frascos foram fechados com rolhas de borracha perfurável do tipo suba-seal estéreis e com o auxílio de uma seringa foi injetado 1mL do gás acetileno (10% da fase gasosa do frasco) e o sistema foi incubado por 1 hora a 30°C. Posteriormente 0,5 mL da fase gasosa do frasco foi injetada, com o auxílio de uma seringa, em cromatógrafo de gás com ionização por chama, marca Perkin 54 Elmer modelo F11, que permitiu a quantificação do etileno formado pela redução do acetileno.

5.3.7 Determinação de proteínas totais

A quantificação de proteínas totais foi realizada utilizando-se o método de Bradford [Bradford, 1976].

Quando a leitura de ARA foi realizada em meio semi-sólido, realizou-se a quantificação de proteínas baseada na técnica descrita por Lowry [Lowry et al., 1951] de acordo com o protocolo descrito por Guedes [Guedes et al., 2007]. Em tubos de ensaio contendo

100 μL da amostra de cultura bacteriana em meio semi-sólido homogeneizada, foram adicionados 400 μL de água destilada estéril e 500 μL de NaOH 1 M. A mistura permaneceu aquecida por 5 minutos a 95°C , para extração das proteínas totais e desnaturação do ágar. Posteriormente foram adicionados 2,5 mL do reagente de Lowry pH 7, e o sistema foi agitado e incubado por 10 minutos no escuro. Em seguida foram adicionados 500 μL do reagente de Folin-Ciocalteu, 1 M, pH 1, e o sistema foi incubado no escuro por 30 minutos. A absorbância foi determinada a 550 nm em espectrofotômetro da marca Bel Photonics modelo SP2100. A concentração de proteína foi calculada usando a curva padrão obtida pelos valores de absorbância de amostras com quantidades conhecidas de BSA (Albumina de Soro Bovino) comercial. A atividade de nitrogenase foi expressa em nmol de etileno por mL de atmosfera. Quando é considerada a concentração de proteínas totais e o tempo de incubação, tem-se então o valor de nmol de etileno por mg ou μg de proteínas totais por unidade de tempo, quantificando desta maneira a FBN dos mutantes comparados com a estirpe padrão PAL5.

5.3.8 Dosagem de aminoácidos livres por HPLC

A dosagem de aminoácidos livres foi realizada de acordo com a técnica de cromatografia líquida de alto desempenho (HPLC) com uma coluna de fase reversa de carbono 18 (22 cm x 4,6 mm) conforme descrito por Hashimoto e colaboradores [Hashimoto et al., 1992].

Para passarem pela coluna, todas as amostras foram previamente desproteinizadas por precipitação com ácido tricloroacético (TCA) em concentração final de 5% seguida de centrifugação. O ácido foi retirado pela adição de um volume de dietil-éter saturado com água seguido de centrifugação. O dietil-éter foi retirado com uma bomba de vácuo. O éter residual foi evaporado a 50°C por aproximadamente 5 min. Os aminoácidos foram separados em um gradiente de acetonitrila. Foi acrescentado tetrahydrofurano aos tampões para aumentar a resolução da separação. Os aminoácidos foram derivatizados com o-ftal-dialdeído e N-tert-butil-oxi-carbonil-L-cisteína no momento da injeção na coluna. A derivatização permite que os isômeros dos aminoácidos sejam eluídos separadamente

e que estes sejam detectados por fluorescência, excitando em 344 nm e absorvendo em 443 nm.

Parâmetros da cromatografia - Volume de injeção: 20 μ L - Fluxo: 0,800 mL/min - Pressão máxima: 200 barr (a pressão normal varia entre 100 e 120) - Fluorescência: Ex = 344, Em = 443

Soluções

- Preparar apenas antes da injeção - 30 μ L de tampão borato pH 9, 23 μ L de água Milli-Q, 5 μ L de L-HCA 50 μ M, 2 μ L de amostra, 20 μ L de derivatizante. Nos casos de adição de padrões, adicionar 5 μ L a 50 μ M, descontar este valor do volume da água.
- Derivatizante - 1 mg de o-phtalaldehide, 1 mg de boc-cys-OH, 100 μ L de metanol (HPLC grade).
- Soluções estoque 1 - NaOH 1M, 8 g de NaOH, 200 mL de água Milli-Q.
- Soluções estoque 2 - Ácido acético 0,2 M, 11,5 mL de ác. acético glacial (17,4 M), água Milli-Q 1 L q.s.p.
- Tampão acetato - Adicionar ~ 195 mL de NaOH 1 M a 1 L de ácido acético 0,2 M. Completar com água Milli-Q até 1995 mL - Ajustar pH 6,2 com ácido acético glacial.
- Tampão borato - Ácido bórico 0,4 M; ajustar pH 9,0 com NaOH.

Tampões de corrida

- Tampão A: 905 mL tampão acetato 0,1 M pH 6,2, 65 mL acetonitrila* 30 mL THF
- Tampão B: 500 mL tampão acetato 0,1 M pH 6,2, 470 mL acetonitrila* 30 mL THF

5.3.9 Varredura dos mutantes auxotróficos para asparagina

Uma colônia de cada mutante aleatório foi inoculada em meio LGI-P sólido com 10 mM de sulfato de amônio ou 10 mM de asparagina como fonte de nitrogênio contendo o antibiótico canamicina (200 $\mu\text{g.mL}^{-1}$) e cultivadas em placa a 30°C por 6 dias. Após o crescimento, as colônias que cresceram no meio com asparagina mas não cresceram no meio com sulfato amônio foram selecionadas para realização dos experimentos posteriores.

Capítulo 6

Visão geral e perspectivas

No artigo onde foi publicado o genoma de *G. diazotrophicus* os autores sugerem que esta bactéria poderia ser um organismo de vida livre que adquiriu certos genes que a permitem colonizar plantas. Essa sugestão foi baseada no fato de que os genes do metabolismo principal de *G. diazotrophicus* são similares a de outras bactérias geneticamente próximas a elas, como *Acidiphilium cryptum* JF-5, *Gluconobacter oxydans* 621H e *Granulibacter bethesdensis* CGDNIH. Uma visão geral desta tese, englobando os três temas abordados, aponta resultados nesta mesma direção.

Os resultados tanto a análise dos transportadores, que indica de a *G. diazotrophicus* possui um repertório de transportadores mais compatível com organismos de vida livre, quanto a análise das sequências das enzimas de detoxificação de ROS e sua capacidade de fixar nitrogênio mesmo na presença de altas concentrações de O_2 , e a análise de sua via para a produção de asparagina indicam que embora a *G. diazotrophicus* seja um organismo extremamente bem adaptado ao crescimento em associação com plantas (no caso mais estudado a cana-de-açúcar), ela possui um grande repertório de genes e características típicas de organismos de vida livre (esquema na Figura 6.1). Isso a torna um organismo extremamente interessante do ponto de vista biológico, com características bem distintas dos fixadores de nitrogênio mais bem estudados.

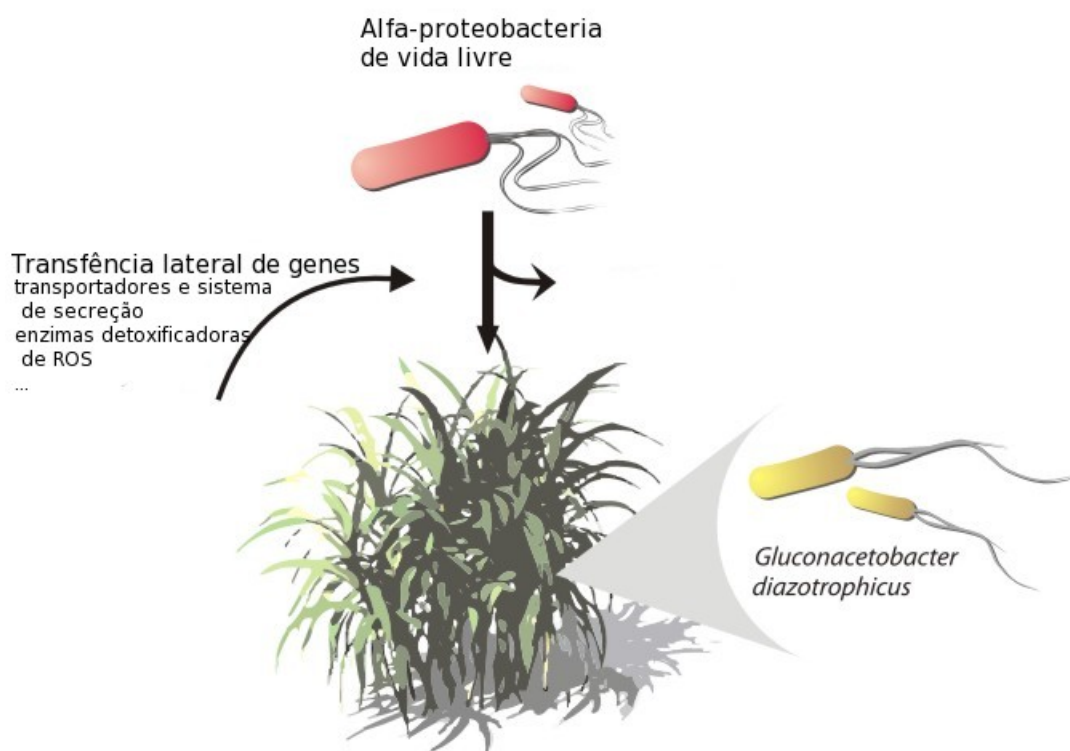


Figura 6.1: Esquema mostrando como a obtenção de determinados genes conferiu vantagens à *G. diazotrophicus* e a permitiu colonizar um novo nicho ecológico.

A disponibilização de seu genoma abriu uma nova era nos estudos com essa bactéria e permitiu que os pesquisadores da área finalmente ligassem o conhecimento biológico da bactéria a seu repertório genético. Essa tese é um exemplo disso: a combinação de análises “in silico”, associadas com experimentos bioquímicos permitiu esclarecer aspectos importantes do metabolismo desta bactéria. A continuidade dos projetos iniciados aqui e de outros projetos de genômica funcional desta bactéria certamente trarão novos “insights” sobre a fisiologia deste organismo e permitirão que o conhecimento gerado com o seu sequenciamento possa por fim ser aplicado em campo.

Referências Bibliográficas

- [Abanda-Nkpwatt et al., 2006] Abanda-Nkpwatt D, Musch M, Tschiersch J, Boettner M, Schwab W (2006) Molecular interaction between *Methylobacterium extorquens* and seedlings: growth promotion, methanol consumption, and localization of the methanol emission site. *J Exp Bot* 57: 4025-4032.
- [Altschul et al., 1990] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- [Altschul et al., 1998] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1998) Gapped BLAST and PSI-BLAST: a new generation of protein database search program. *Nucleic Acids Res* 25: 3389-3402.
- [Alvarez e Drets, 1995] Alvarez B, Drets GM (1995) Metabolic characterization of *Acetobacter diazotrophicus*. *Can J Microbiol* 41: 918-924.
- [Ambrogelly et al., 2007] Ambrogelly A, Palioura S, Söll D (2007) Natural expansion of the genetic code. *Nature Chemical Biology* 3: 29-35.
- [Apel e Hirt, 2004] Apel K, Hirt H (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* 55: 373-399.
- [Ashbolt e Inkerman, 1990] Ashbolt PA, Inkerman NJ (1990) Acetic Acid Bacterial Biota of the Pink Sugar Cane Mealybug, *Saccharococcus sacchari*, and its environs. *Applied and Environmental Microbiology* 56: 707-712.
- [Attwood et al., 1991] Attwood MM, van Dijken JP, and Pronk JT (1991) Glucose metabolism and gluconic acid production by *Acetobacter diazotrophicus*. *J Ferment Bioeng* 72: 101-105.
- [Aydi et al., 2004] Aydi S, Drevon JJ, Abdelly C (2004) Effect of salinity on root-nodule conductance to the oxygen diffusion in the *Medicago truncatula* – *Sinorhizobium meliloti* symbiosis. *Plant Physiol Biochem* 42: 833-840.
- [Baldani et al., 1983] Baldani VLD, Baldani JI, Döbereiner J (1983) Effects of *Azospirillum* inoculation on root infection and nitrogen incorporation in wheat. *Can J Microbiol* 29: 924-929.

- [Baldani et al., 1997] Baldani JI, Caruso L, Baldani VLD, Goi SR, Döbereiner J (1997) Recent advances in BNF with non-legume plants. *Soil Biology & Biochemistry* 29: 911-922.
- [Baldani et al., 1999] Baldani JI, Azevedo MS, Reis VM, Teixeira KRS, Olivares FL, Goi SR, Baldani VLD, Döbereiner J (1999) Fixação biológica de nitrogênio em gramíneas: avanços e aplicações. Em: Siqueira JO, Moreira FMS, Lopes AS, Guilherme LRG, Faquin V, Furtini Neto AE, Carvalho JG (eds) Inter-relação fertilidade, biologia do solo e nutrição de plantas. p. 621-666. SBCS/UFLA, Viçosa.
- [Baldani e Baldani, 2005] Baldani JI, Baldani VL (2005) History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience. *An Acad Bras Cienc* 77: 549-579.
- [Bateman et al., 2004] Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme D.J, Yeats C, Eddy SR (2004) The Pfam protein families database. *Nucleic Acids Res* 32: 138-141.
- [Bashan e Levanony, 1990] Bashan Y, Levanony H (1990) Current status of Azospirillum as a challenge for agriculture. *Canadian Journal of Microbiology* 36: 591-608.
- [Blanco et al., 2005] Blanco Y, Blanch M, Piñón D, Legaz ME, Vicente C (2005) Antagonism of *Gluconacetobacter diazotrophicus* (a sugarcane endosymbiont) against *Xanthomonas albilineans* (pathogen) studied in alginate-immobilized sugarcane stalk tissues. *Journal of Bioscience and Bioengineering* 99: 366-371.
- [Blanvillain et al., 2007] Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M (2007) Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS One* 2: 224.
- [Bradford, 1976] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal Biochem* 72: 248-254.
- [Berman et al, 2000] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, (2000) The protein data bank. *Nucleic Acids Res* 28: 235-242.
- [Bertalan et al, 2009] Bertalan M, Albano R, de Pádua V, Rouws L, Rojas C, Hemerly A, Teixeira K, Schwab S, Araujo J, Oliveira A, França L, Magalhães V, Alquéres S, Cardoso A, Almeida W, Loureiro MM, Nogueira E, Cidade D, Oliveira D, Simão T, Macedo J, Valadão A, Dreschel M, Freitas F, Vidal M, Guedes H, Rodrigues E, Meneses C, Brioso P, Pozzer L, Figueiredo D, Montano H, Junior J, de Souza Filho G, Martin Quintana Flores V, Ferreira B, Branco A, Gonzalez P, Guillobel H,

- Lemos M, Seibel L, Macedo J, Alves-Ferreira M, Sachetto-Martins G, Coelho A, Santos E, Amaral G, Neves A, Pacheco AB, Carvalho D, Lery L, Bisch P, Rössle SC, Urményi T, Rael Pereira A, Silva R, Rondinelli E, von Krüger W, Martins O, Baldani JI, Ferreira PC (2009) Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. *BMC Genomics* 10: 450.
- [Boddey et al., 1991] Boddey RM, Urquiaga S, Reis V, Döbereiner J (1991) Biological nitrogen fixation associated with sugar cane. *Plant and Soil* 137: 111-117.
- [Boeckmann et al., 2003] Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin M.J, Michoud K, O'Donovan C, Phan I, Pilbout S, Schneider M (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* 31: 365-370.
- [Bonfield et al., 1995] Bonfield JK, Smith KF, Staden R (1995) A new DNA sequence assembly program. *Nucleic Acids Res* 23: 4992-4999.
- [Boveris e Chance, 1973] Boveris A, Chance B (1973) The mitochondrial generation of hydrogen peroxide. General properties and effect of hyperbaric oxygen. *Biochem J* 134: 707-716.
- [Caballero-Mellado et al., 1995] Caballero-Mellado J, Fuentes-Ramírez LE, Reis VM, Martínez-Romero E (1995) Genetic Structure of *Acetobacter diazotrophicus* Populations and Identification of a New Genetically Distant Group. *Applied and Environmental Microbiology* 61: 3008-3012.
- [Caballero-Mellado et al., 1999] Caballero-Mellado J, Jiménez-Salgado T, Tapia-Hernández A, Wang E T, Martínez-Romero E and Fuentes-Ramírez L E (1999) Polyphasic taxonomy of nitrogen-fixing acetic bacteria isolated from the rhizosphere of coffee plants. Em: Martínez E, Hernández G (eds) Highlights on nitrogen fixation research, p. 271-274. Kluwer Academic Publisher, New York.
- [Calendar e Berg 1967] Calendar R, Berg P (1967) D-Tyrosyl RNA: formation, hydrolysis and utilization for protein synthesis. *J Mol Biol* 26: 39-54.
- [Capone et al., 1994] Capone DG, Ferrier MD, Carpenter EJ (1994) Amino acid cycling in colonies of the marine planktonic cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* 60: 3989-3995.
- [Cardenas et al., 2008] Cardenas, L., Martinez, A., Sanchez, F., and Quinto, C. (2008). Fast, transient and specific intracellular ROS changes in living root hair cells responding to Nod factors (NFs). *Plant J* 56: 802-813.
- [Carol e Dolan, 2006] Carol RJ, Dolan L (2006) The role of reactive oxygen species in cell growth: lessons from root hairs. *J Exp Bot* 57: 1829-1834.

- [Carraro e Kitajima, 2002] Carraro DM, Kitajima JP (2002) Seqüenciamento e bioinformática de genomas bacterianos. *Biotecnologia Ciência & Desenvolvimento* 28: 16-20.
- [Carver et al., 2008] Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream MA (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24: 2672-2676.
- [Castro-Sowinski et al., 2007] Castro-Sowinski S, Matan O, Bonafede P, Okon Y (2007) A thioredoxin of *Sinorhizobium meliloti* CE52G is required for melanin production and symbiotic nitrogen fixation. *Mol Plant Microbe Interact.* 20: 986-993.
- [Cavalcante e Döbereiner, 1988] Cavalcante VA, Döbereiner J (1988) A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. *Plant and Soil* 108: 23-31.
- [Cedar e Schwartz, 1968] Cedar H, Schwartz JH (1968) Production of L-Asparaginase II by *Escherichia coli*. *J Bacteriol* 96: 2043-2048.
- [Chang et al, 2009] Chang C, Damiani I, Puppo A, Frendo P (2009) Redox changes during the Legume - *Rhizobium* symbiosis. *Molecular Plant* 2: 370-377.
- [Chen et al, 1990] Chen MW, Jahns D, Schön A, O'Neill GP, Söll D (1990) Purification and characterization of *Chlamydomonas reinhardtii* chloroplast glutamyl-tRNA synthetase, a natural misacylating enzyme. *Journal of Biological Chemistry* 265: 4054-4057.
- [Chimento et al., 2003] Chimento DP, Kadner RJ, Wiener MC (2003) The *Escherichia coli* outer membrane cobalamin transporter BtuB: structural analysis of calcium and substrate binding, and identification of orthologous transporters by sequence/structure conservation. *J Mol Biol* 332: 999-1014.
- [Chomczynski e Sacchi, 1987] Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162: 156-159.
- [Chuawong e Hendrickson, 2006] Chuawong P, Hendrickson TL (2006) The non-discriminating aspartyl-tRNA synthetase from *Helicobacter pylori*: Anticodon-binding domain mutations that impact tRNA specificity and heterologous toxicity. *Biochemistry* 45: 8079-8087.
- [Cirousse et al., 1996] Cirousse C, Bournoville R, Bonnemain JL (1996) Water deficit-induced changes in concentrations in proline and some other amino acids in the phloem sap of alfalfa. *Plant Physiol* 111: 109-113.
- [Chollet et al, 1996] Chollet R, Vidal J, O'Leary MH (1996) Phosphoenolpyruvate carboxylase: A Ubiquitous, Highly Regulated Enzyme in Plants. *Annu Rev Plant Physiol Plant Mol Biol* 47: 273-298.

- [Coschrane et al., 2006] Cochrane G, Aldebert P, Althorpe N, Andersson M, Baker W, Baldwin A, Bates K, Bhattacharyya S, Browne P, van den Broek A, Castro M, Duggan K, Eberhardt R, Faruque N, Gamble J, Kanz C, Kulikova T, Lee C, Leinonen R, Lin Q, Lombard V, Lopez R, McHale M, McWilliam H, Mukherjee G, Nardone F, Pastor MP, Sobhany S, Stoehr P, Tzouvara K, Vaughan R, Wu D, Zhu W, Apweiler R (2006) EMBL Nucleotide Sequence Database: developments in 2005. *Nucleic Acids Res* 34: 10-15.
- [Cote e Ahlgran, 1993] Cote GL, Ahlgran JA (1993) Metabolism in microorganism Part I. Levan and levansucrase. Em: M Susuki, Chatterton NJ (eds) *Science and Technology of Fructans*, p. 141-168. CRC Press, Boca Raton.
- [Crick, 1970] Crick F (1970) Central Dogma of Molecular Biology. *Nature* 227: 561-563.
- [Curnow et al., 1996] Curnow AW, Ibba M, Söll D (1996) tRNA-dependent asparagine formation. *Nature* 382: 589-590.
- [Curnow et al., 1997] Curnow AW, Honk KW, Yuan R, KIM SI, Martins OB, Winkler W, Henkin TM, Söll D (1997). *Glu* – *tRNA^{Gln}* amidotransferase: A novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proceedings of the National Academy of Sciences USA* 94: 11819-11826.
- [Curnow et al., 1998] Curnow AW, Tumbula DL, Pelaschier JT, Min B, Söll D (1998). *Glutamyl* – *tRNA^{Gln}* amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proceedings of the National Academy of Sciences USA* 95: 12838-12843.
- [Dalton e Postgate, 1968] Dalton H, Postgate JR (1968) Effect of oxygen on growth of *Azotobacter chroococcum* in batch and continuous cultures. *J Gen Microbiol* 54: 463-473.
- [Dalton, 1995] Dalton DA (1995) Antioxidant defenses of plants and fungi. Chapman & Hall, New York.
- [Davies e Walker 2007] Davies BW, Walker GC (2007) Identification of novel *Sinorhizobium meliloti* mutants compromised for oxidative stress protection and symbiosis. *J Bacteriol* 189: 2110-2113.
- [Del Castillo et al., 1994] Del Castillo LD, Hunt S, Layzell DB (1994) The role of oxygen in the regulation of nitrogenase activity in drought stressed soybean nodules. *Plant Physiol* 106: 949-955.
- [Delcher et al., 1999] Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* 27: 4636-4641.

- [De Ley et al., 1984] De Ley J, Gillis M and Swings J (1984) Family VI. *Acetobacteraceae*. Em: Krieg NR, Holt JG (eds) Bergey's Manual of Systematic Bacteriology 1: 267-278. Williams and Wilkins, Baltimore.
- [Detroy et al., 1967] Detroy RW, Witz DF, Wilson PW (1967) Nitrogen fixation by growing cells and cell-free extract of the Bacillaceae. Archives of microbiology 55: 369-381.
- [D'Haeseleer, 2005] D'Haeseleer P (2005) How does gene expression clustering work? Nat Biotechnol 23: 1499-1501.
- [Dixon e Kahn, 2004] Dixon R, Kahn D (2004) Genetic regulation of biological nitrogen fixation. Nature Reviews Microbiology 2: 621-631.
- [Döbereiner, 1992] Döbereiner J (1992) Recent changes in concepts of plant bacteria interactions: endophytic N_2 fixing bacteria. Ciência e Cultura 44: 310-313.
- [Döbereiner et al., 1993] Döbereiner J, Reis VM, Paula MA, Olivares FL (1993) Endophytes diazotrophs in sugar cane, cereals and tuber plants. Em Palacios R, Mora J, Newton WE (eds) New horizons in nitrogen fixation p. 671-676. Kluwer Academic Publishers, Dordrecht.
- [Döbereiner et al., 1995] Döbereiner J, Bandalni VLD, Baldani JI (1995) Como isolar e identificar bactérias diazotróficas em plantas não leguminosas. EMBRAPA – SPI, Brasília. EMBRAPA – CNPAB, Itaguaí-RJ.
- [Dombrecht et al., 2005] Dombrecht B, Heusdens C, Beullens S, Verreth C, Mulkers E, Proost P, Vanderleyden J, Michiels J (2005) Defence of *Rhizobium etli* bacteroids against oxidative stress involves a complexly regulated atypical 2-Cys peroxiredoxin. Mol Microbiol 55: 1207-1221.
- [Dong et al., 1994] Dong Z, Canny MJ, McCully ME, Roboredo MR, Cabadilla CF, Ortega E, Rodes R (1994) A Nitrogen-Fixing Endophyte of Sugarcane Stems (A New Role for the Apoplast). Plant Physiol 105: 1139-1147.
- [Dong et al., 1995] Dong Z, Heydrich M, Bernard K, McCully ME (1995) Further evidence that the N_2 -fixing endophytic bacterium from the intercellular spaces of sugarcane stems is *Acetobacter diazotrophicus*. Applied and Environmental Microbiology 61: 1843-1846.
- [dos Santos et al., 2010] dos Santos MF, de Padua VL, Nogueira EM, Hemerly AS, Domont GB (2010) Proteome of *Gluconacetobacter diazotrophicus* co-cultivated with sugarcane plantlets. Journal of Proteomics 73: 917-31.
- [Downs e Escalante-Semerena, 2000] Downs DM, Escalante-Semerena JC (2000) Impact of Genomics and Genetics on the Elucidation of Bacterial Metabolism. Methods 20: 47-54.

- [Doyle e Doyle, 1987] Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11-15.
- [Eady, 1996] Eady RR (1996) Structure–function relationships of alternative nitrogenases. *Chem Reviews* 96: 3013-3030.
- [Emanuelsson et al., 2007] Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953-971.
- [Falquet et al., 2002] Falquet L, Pagni M, Bucher P, Hulo N, Sigrist C.J, Hofmann K, Bairoch A (2002) The PROSITE database, its status in 2002. *Nucleic Acids Res* 30: 235-238.
- [Felsenstein, 2006] Felsenstein J (2006) Department of Genetics, University of Washington, PHYLIP: the PHYLogeny Inference Package. Disponível em: <http://evolution.genetics.washington.edu/phylip.html>. Acesso em: 10 de maio de 2006.
- [Feng et al., 2005] Feng L, Yuan J, Toogood H, Tumbula-hansen D, Söll D (2005) Aspartyl-tRNA synthetase requires a conserved proline in the anticodon-binding loop for *tRNA^{Asn}* recognition in vivo. *Journal of Biological Chemistry* 280: 20638-20641.
- [Fichant et al., 2009] Fichant G, Basse MJ, Quentin Y (2006) ABCdb: an online resource for ABC transporter repertoires from sequenced archaeal and bacterial genomes. *FEMS Microbiol Lett* 256: 333-339.
- [Flores-Encarnacion et al., 1999] Flores-Encarnacion M, Contreras-Zentella M, Soto-Urzuá L, Aguilar GR, Baca BE, Escamilla JE (1999) The respiratory system and diazotrophic activity of *Acetobacter diazotrophicus* PAL5. *J Bacteriol* 181: 6987-6995.
- [Forman et al., 2010] Forman HJ, Maiorino M, Ursini F (2010) Signaling Functions of Reactive Oxygen Species. *Biochemistry* 49: 835-842.
- [Fuentes-Ramirez et al., 1993] Fuentes-Ramirez LE, Jimenez-Salgado T, Abarca-Ocampo LR, Caballero-Mellado J (1993) *Acetobacter diazotrophicus* a endolactic producing bacterium isolated from sugar cane cultivars of Mexico. *Plant and Soil* 154: 145-150.
- [Fuentes-Ramírez et al., 2001] Fuentes-Ramírez LE, Bustillos-Cristales R, Tapia-Hernández A, Jiménez-Salgado T, Wang ET, Martínez-Romero E, Caballero-Mellado J (2001) Novel nitrogen-fixing acetic acid bacteria, *Gluconacetobacter johannae* sp. nov., and *Gluconacetobacter azotocaptans* sp. nov., associated with coffee plants. *Int J Syst Evol Microbiol* 51: 1305-1314.

- [Gagnon et al., 1996] Gagnon Y, Lacoste L, Champagne N, Lapointe J (1996) Widespread use of the *Glu* – *tRNA^{Gln}* transamidation pathway among bacteria. *Journal of Biological Chemistry* 271: 14856-14863.
- [Gehrig et al., 1998] Gehrig HH, Heute V, Kluge M (1998) Toward a better knowledge of the molecular evolution of phosphoenolpyruvate carboxylase by comparison of partial cDNA sequences. *J Mol Evol* 46: 107-114.
- [Gillis et al., 1989] Gillis M, Kersters K, Hoste B, Janssens D, Kroppenstedt RM, Stephan MP, Teixeira KRS, Döbereiner J, De Ley J (1989) *Acetobacter diazotrophicus* sp. nov., a nitrogen-fixing acetic acid bacterium associated with sugarcane. *International Journal of Systematic Bacteriology* 39: 361-364.
- [Gonzalez et al., 2006] Gonzalez B, Martinez S, Chavez JL, Lee S, Castro NA, Dominguez MA, Gomez S, Contreras ML, Kennedy C, Escamilla JE (2006) Respiratory system of *Gluconacetobacter diazotrophicus* PAL5. Evidence for a cyanide-sensitive cytochrome bb and cyanide-resistant cytochrome ba quinol oxidases. *Biochim Biophys Acta* 1757: 1614-1622.
- [Goran, 1947] Goran M (1947) The Present-Day Significance of Fritz Haber. *American Scientist* 35: 400-403.
- [Gordon et al., 2001] Gordon AJ, Lea PJ, Rosenberg C, Trinchant JC (2001) Nodule formation and function in legumes. Em: Lea PJ, Morot-Gaudry JF (eds) *Plant Nitrogen* p. 101-146. Springer-Verlag, Berlin.
- [Goldberg et al., 1987] Goldberg I, Nadler V, Hochman A (1987) Mechanisms of nitrogenase switch-off by oxygen. *J Bacteriol* 169: 874-879.
- [Gechev et al., 2006] Gechev TS, Van Breusegem F, Stone JM, Denev I, Laloi C (2006). Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *Bioessays* 28 1091-1101.
- [Guedes et al., 2007] Guedes HV, Perin L, Reis VM, Baldani JI, Teixeira KRS (2007) Quantificação de proteínas totais de bactérias diazotróficas crescidas em meio de cultivo semi-sólido. *Comunicado técnico* 95, ISSN 1517-8862.
- [Guedes, 2010] Guedes HV (2010) Identificação e caracterização de genes envolvidos no processo de fixação biológica do nitrogênio em *Gluconacetobacter diazotrophicus* PAL5: ênfase em um gene do tipo histidina quinase. Rio de Janeiro, 2010. Tese (Doutorado em Biotecnologia Vegetal) – Centro de Ciências da Saúde (CCS), Universidade Federal do Rio de Janeiro.
- [Halliwell e Gutteridge, 1999] Halliwell B, Gutteridge JMC (1999) *Free Radicals in Biology and Medicine*. Oxford University Press, New York.

- [Harger et al., 1998] Harger C, Skupski M, Bingham J, Farmer A, Hoisie S, Hraber P, Kiphart D, Krakowski L, McLeod M, Schwertfeger J, Seluja G, Siepel A, Singh G, Stamper D, Steadman P, Thayer N, Thompson R, Wargo P, Waugh M, Zhuang J.J, Schad PA (1998) The Genome Sequence DataBase (GSDB): improving data quality and data access. *Nucleic Acids Res* 26: 21-26.
- [Harrison et al., 2005] Harrison J, Jamet A, Muglia CI, Van de Sype G, Aguilar OM, Puppo A, Frendo P (2005) Glutathione plays a fundamental role in growth and symbiotic capacity of *Sinorhizobium meliloti*. *J Bacteriol* 187: 168-174.
- [Hashimoto et al., 1992] Hashimoto A, Nishikawa T, Oka T, Takahashi K, Hayashi T (1992) Determination of free amino acid enantiomers in rat brain and serum by high-performance liquid chromatography after derivatization with N-tert.-butyloxycarbonyl-L-cysteine and o-phthalaldehyde. *J Chromatogr* 582: 41-48.
- [Hassett et al., 1987] Hassett DJ, Britigan BE, Svendsen T, Rosen GM, Cohen MS (1987) Bacteria form intracellular free radicals in response to paraquat and streptonigrin. Demonstration of the potency of hydroxyl radical. *J Biol Chem* 262: 13404-13408.
- [Hernández et al., 1995] Hernández L, Arrieta J, Menéndez C, Vazquez R, Coego A, Suárez V, Selman G, Petit-Glatron M F and Chambert R (1995) Isolation and enzymic properties levansucrase secreted by *Acetobacter diazotrophicus* SRT4, a bacterium associated with sugar cane. *Biochem J* 309: 113-118.
- [Hernández et al., 2000] Hernández L, Sotolongo M, Rosabal Y, Menéndez C, Ramírez R, Caballero-Mellado J and Arrieta J (2000) Structural levansucrase gene (*lsdA*) constitutes a functional locus conserved in the species *Gluconacetobacter diazotrophicus*. *Arch Microbiol* 174: 111-119.
- [Hua et al., 1982] Hua SST, Tsai VY, Lichens GM, Noma AT (1982) Accumulation of amino acids in *Rhizobium* sp. Strain WR1001 in Response to Sodium Chloride Salinity. *Applied and Environmental Microbiology* 44: 135-140.
- [Huergo et al., 2007] Huergo LF, Merrick M, Pedrosa FO, Chubatsu LS, Araujo LM, Souza EM (2007) Ternary complex formation between AmtB, GlnZ and the nitrogenase regulatory enzyme DraG reveals a novel facet of nitrogen regulation in bacteria. *Mol Microbiol* 66: 1523-1535.
- [Humbert e Simoni, 1980] Humbert R, Simoni RD (1980) Genetic and biochemical studies demonstrating a second gene coding for asparagine synthetase in *Escherichia coli*. *J Bacteriol* 142: 212-220.
- [Ibba e Söll, 2000] Ibba M, Söll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69: 617-650.

-
- [Ibba et al., 2000] Ibba M, Becker HD, Stathouopoulos C, Tumbula DL, Söll D (2000) The adaptor hypothesis revisited. *Trends in Biochemistry Science* 25: 311-316.
- [James, 2000] James EK (2000) Nitrogen fixation in endophytic and associative symbiosis. *Field Crops Research* 65: 197-209.
- [Jamet et al., 2003] Jamet A, Sigaud S, Van de Sype G, Puppo A, Herouart D (2003) Expression of the bacterial catalase genes during *Sinorhizobium meliloti* – *Medicago sativa* symbiosis and their crucial role during the infection process. *Mol Plant Microbe Interact.* 16: 217-225.
- [Jamet et al., 2007] Jamet A, Mandon K, Puppo A, Herouart D (2007) H_2O_2 is required for optimal establishment of the *Medicago sativa* - *Sinorhizobium meliloti* symbiosis. *J Bacteriol* 189: 8741-8745.
- [Jiménez-Salgado et al., 1997] Jiménez-Salgado T, Fuentes-Ramírez LE, Tapia-Hernández A, Mascarua-Esparza MA, Martínez-Romero E, Caballero-Mellado J (1997) *Coffea arabica* a new host plant for *Gluconacetobacter diazotrophicus* and isolation of other nitrogen-fixing acetobacteria. *Applied and Environmental Microbiology* 63: 3676-3683.
- [Jones, 2006] Jones DT (2006) Threader – Protein fold recognition by optimal protein sequence threading. Disponível em <http://bioinf.cs.ucl.ac.uk/threader/>. Acesso em 10 de maio de 2006.
- [Kanamori et al., 1988] Kanamori K, Weiss RL, Roberts JD (1988) Glutamate biosynthesis in *Bacillus azotofixans*. SN NMR and enzymatic studies. *Journal of Biological Chemistry* 263: 2817-2823.
- [Kaneko et al., 2003] Kaneko T, Suzuki T, Kapushoc ST, Rubio MA, Ghazvini J, Watanabe K, Simpson L, Suzuki T (2003) Wobble modification differences and subcellular localization of tRNAs in *Leishmania tarentolae*: implication for tRNA sorting mechanism. *EMBO J* 22: 657-667
- [Kapulnik et al., 1985] Kapulnik Y, Gafny R, Okon Y (1985) Effect of *Azospirillum spp.* inoculation on root development and NO_3 -uptake in wheat (*Triticum aestivum*, cv. Miriam) in hydroponic systems. *Canadian Journal of Botany* 63: 627-631.
- [Kim e Rees, 1994] Kim J, Rees DC (1994) Nitrogenase and biological nitrogen fixation. *Biochemistry* 33: 389-397.
- [Koebnik, 2005] Koebnik R (2005) TonB-dependent trans-envelope signalling: the exception or the rule? *Trends Microbiol* 13: 343-347.
- [Koonin et al., 1996] Koonin EV, Mushegian AR, Rudd KE (1996) Sequencing and analysis of bacterial genomes. *Curr Biology* 6: 404-416.
-

- [Kowaltowski et al., 2009] Kowaltowski AJ, de Souza-Pinto NC, Castilho RF, Vercesi AE (2009) Mitochondria and reactive oxygen species. *Free Radical Biology & Medicine* 47: 333-343.
- [Krause et al., 2006] Krause A, Ramakumar A, Bartels D, Battistoni F, Bekel T, Boch J, Böhm M, Friedrich F, Hurek T, Krause L, Linke B, McHardy AC, Sarkar A, Schneiker S, Syed AA, Thauer R, Vorhölter FJ, Weidner S, Purhler A, Reinhold-Hurek B, Kaiser O, Goesmann A (2006) Complete genome of the mutualistic N_2 -fixing grass endophyte *Azoarcus* sp. strain BH72. *Nat Biotech* 24: 1385-1391.
- [Krogh et al., 2010] Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567-680.
- [Lapointe et al., 1986] Lapointe J, Duplain L, Proulx M (1986) A single glutamyl-tRNA synthetase aminoacylates *tRNA^{Glu}* and *tRNA^{Gln}* in *Bacillus subtilis* and efficiently misacylates *Escherichia coli* *ttRNA₁^{Gln}* in vitro. *Journal of Bacteriology* 165: 88-93.
- [Lawson e Smith, 2002] Lawson DM, Smith BE (2002) Metal Ions in Biological Systems. Em: Sigel A, Sigel H (eds) 39: 75-119. Marcel Dekker, New York.
- [Lee et al., 2004] Lee S, Flores-Encarnacion M, Contreras-Zentella M, Garcia-Flores L, Escamilla JE, Kennedy C (2004). Indole-3-Acetic Acid Biosynthesis Is Deficient in *Gluconacetobacter diazotrophicus* Strains with Mutations in Cytochrome c Biogenesis Genes. *J Bacteriol* 186: 5384-5391.
- [Lemos, 2004] Lemos M (2004) Workflow para bioinformática. Tese de Doutorado em Ciências da Computação. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- [Lemos et al., 2006] Lemos M, Seibel LFB, Casanova MA (2003) Sistema de Anotações em Biosseqüências. Monografia da ciência da computação número 04/03, Departamento de Informática. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.
- [Lepiniec et al., 1994] Lepiniec L, Vidal J, Chollet R, Gadal P, Créatin C (1994) Phosphoenolpyruvate carboxylase: structure, regulation and evolution. *Plant Sci* 99: 111-124.
- [Lery et al., 2008] Lery LM, von Kruger WM, Viana FC, Teixeira KR, Bisch PM (2008) A comparative proteomic analysis of *Gluconacetobacter diazotrophicus* PAL5 at exponential and stationary phases of cultures in the presence of high and low levels of inorganic nitrogen compound. *Biochim Biophys Acta* 1784: 1578-1589.
- [Letovsky et al., 2006] Letovsky SI, Cottingham RW, Porter CJ, Li PW (2006) GDB: the Human Genome Database. *Nucleic Acids Res* 26: 94-99.

- [Lee et al., 2000] Lee AK, Detweiler CS, Falkow S (2000) OmpR regulates the two-component system SsrA-SsrB in *Salmonella* pathogenicity island 2. *Journal Bacteriology* 182: 771-781.
- [Li e MacRae, 1991] Li R, MacRae IC (1991) Specific association of diazotrophic *Acetobacters* with sugarcane. *Soil Biol Biochem* 23: 999-1002.
- [Linu et al., 2009] Linu MS, Stephen J, Jisha MS (2009) Phosphate solubilizing *Glucacetobacter sp.*, *Burkholderia sp.* and their potential interaction with Cowpea (*Vigna unguiculata L.*). *International Journal of Agricultural Research* 4: 79-87.
- [Liolios et al., 2009] Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC (2009) The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *NAR Epub*.
- [Livak e Schmittgen, 2001] Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25: 402-408.
- [Loganathan et al., 1999] Loganathan P, Sunitha R, Parida AK, Nair S (1999) Isolation and characterisation of two genetically distant groups of *Acetobacter diazotrophicus* from new host plant (*Eleusine coracana L.*). *J Appl Bacteriol* 86: 1053-1058.
- [Loiret et al., 2004] Loiret FG, Ortega E, Kleiner D, Ortega-Rodés P, Rodés R, Dong Z (2004) A putative new endophytic nitrogen-fixing bacterium *Pantoea sp.* form sugarcane. *J Appl Microbiol* 97: 504-511.
- [Loiret et al., 2009] Loiret FG, Grimm B, Hajirezaei MR, Kleiner D, Ortega E (2009) Inoculation of sugarcane with *Pantoea sp.* increases amino acid contents in shoot tissues; serine, alanine, glutamine and asparagine permit concomitantly ammonium excretion and nitrogenase activity of the bacterium. *Journal of Plant Physiology* 166: 1152-1161.
- [Lowe et al., 1997] Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955-964.
- [Lowry et al., 1951] Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193: 265-275.
- [Ludwig, 1980] Ludwig RA (1980) Regulation of *Rhizobium* nitrogen fixation by the unadenylylated glutamine synthetase I system. *Proc Natl Acad Sci USA* 77: 5817-5821.
- [Martin e Mohn, 1999] Martin VJ, Mohn WW (1999) An alternative inverse PCR (iPCR) method to amplify DNA sequences flanking Tn5 transposon insertions. *J Microbiol Meth* 35: 163-166.

- [Mazzafera e Gonçalves, 1999] Mazzafera P, Gonçalves KV (1999) Nitrogen compounds in the xylem sap of coffee. *Phytochemistry* 50: 383-386.
- [McClain, 1993] McClain WH (1993) Rules that govern tRNA identity in protein synthesis. *Journal of Molecular Biology* 234: 257-280.
- [McCully, 2001] McCully ME (2001) Niches for endophytes in crop plants: a plant biologist view. *Aust J Plant Physiol* 28: 983-990.
- [Merrick, 1992] Merrick MJ (1992) Regulation of Nitrogen Fixation Genes in Free-living and Symbiotic Bacteria. Em: Stacey G, Burris RH, Evans HJ (eds) *Biological Nitrogen Fixation* p. 835-876. London: Chapman & Hall, New York.
- [Merrick e Edwards, 1995] Merrick MJ, Edwards RA (1995) Nitrogen control in bacteria. *Microbiology Reviews* 59: 604-622.
- [Min et al., 2002] Min B, Pelaschier JT, Graham DE, Hansen DT, Söll S (2002) Transfer RNA-dependent amino acid biosynthesis: An essential route to asparagine formation. *PNAS* 99: 2678-2683.
- [Moller et al., 2007] Moller IM, Jensen PE, Hansson A (2007). Oxidative modifications to cellular components in plants. *Annu Rev Plant Biol* 58: 459-481.
- [Moszer, 1998] Moszer I (1998) The complete genome of *Bacillus subtilis*: from sequence annotation to data management and analysis. *FEBS Lett* 430: 28-36.
- [Muñoz-Rojas e Caballero-Mellado, 2003] Muñoz-Rojas J, Caballero-Mellado J (2003) Population Dynamics of *Gluconacetobacter diazotrophicus* in Sugarcane Cultivars and Its Effect on Plant Growth. *Microbial Ecology* 46: 454-464.
- [Muthukumarasamy et al., 2002] Muthukumarasamy R, Revathi G, Seshadri S, Lakshminarasimhan C (2002) *Gluconacetobacter diazotrophicus* (syn. *Acetobacter diazotrophicus*), a promising diazotrophic endophyte in tropics. *Current Science* 83: 137-145.
- [Muthukumarasamy et al., 2006] Muthukumarasamy R, Govindarajan M, Vadivelu M, Revathi G (2006) N-fertilizer saving by the inoculation of *Gluconacetobacter diazotrophicus* and *Herbaspirillum sp.* in micropropagated sugarcane plants. *Microbiol Res* 161: 238-245.
- [NCBI, 2006a] NCBI – National Center for Biotechnology Information. Open Reading Frame Finder. Disponível em: <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>. Acesso em 10 de maio de 2006a.
- [NCBI, 2006b] NCBI – National Center for Biotechnology Information. The Blast Databases. Disponível em: <ftp://ftp.ncbi.nih.gov/blast/db/>. Acesso em 10 de maio de 2006b.

- [Oliveira et al., 2002] Oliveira ALM, Urquiaga S, Döbereiner J, Baldani JI (2002) The effect of inoculating endophytic N₂-fixing bacteria on micropropagated sugarcane plants. *Plant and Soil* 242: 205-215.
- [Oliveira et al., 2006] Oliveira ALM, Canuto EL, Urquiaga S, Reis VM, Baldani JI (2006) Yield of micropropagated sugarcane varieties in different soil types following inoculation with diazotrophic bacteria. *Plant and Soil* 284: 23-32.
- [Overmyer et al., 2003] Overmyer K, Brosche M, Kangasjarvi J (2003) Reactive oxygen species and hormonal control of cell death. *Trends Plant Sci* 8: 335-342.
- [Palmgren e Axelsen, 1998] Palmgren MG, Axelsen KB (1998) Evolution of P-type ATPases. *Bioch Biophys Acta* 1365: 37-54.
- [Pan e Vessey, 2001] Pan B, Vessey JK (2001) Response of the endophytic diazotroph *Gluconacetobacter diazotrophicus* on solid media to changes in atmospheric partial O₂ pressure. *Appl Environ Microbiol* 67: 4694-4700.
- [Paula et al., 1991] Paula MA, Reis VM, Döbereiner J (1991) Interactions of *Glomus clarum* with *Acetobacter diazotrophicus* in infection of sweet potato (*Ipomoea batata*), sugarcane (*Saccharum sp.*), and sweet sorghum (*Sorghum vulgare*). *Biology and Fertility of Soils* 11: 111-115.
- [Paula et al., 1992] Paula MA, Urquiaga S, Siqueira JO, Döbereiner J (1992) Synergistic effects of vesicular-arbuscular mycorrhizal fungi and diazotrophic bacteria on nutrition and growth of sweet potato (*Ipomoea batatas*). *Biology and Fertility of Soils* 14: 61-66.
- [Paulsen et al., 1998] Paulsen IT, Sliwinski MK, Saier MHJr. (1998) Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *J Mol Biol* 277: 573-592.
- [Paulsen et al., 1996] Paulsen IT, Brown MH, Skurray RA (1996) Proton-dependent multidrug efflux systems. *Microbiol Rev* 60: 575-608.
- [Peleg-Grossman et al., 2007] Peleg-Grossman S, Volpin H, Levine A (2007) Root hair curling and Rhizobium infection in *Medicago truncatula* are mediated by phosphatidylinositide-regulated endocytosis and reactive oxygen species. *J Exp Bot* 58: 1637-1649.
- [Pfaffl, 2001] Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29: e45.
- [Piñón et al., 2002] Piñón D, Casas M, Blanch M, Fontaniella B, Blanco Y, Vicente C, Solas MT, Legaz ME (2002) *Gluconacetobacter diazotrophicus*, a sugarcane endosymbiont, produces a bacteriocin against *Xanthomonas albilineans*, a sugarcane pathogen. *Res Microbiol* 153: 345-351.

- [Pitzschke et al., 2006] Pitzschke A, Forzani C, Hirt H (2006). Reactive oxygen species signaling in plants. *Antioxid Redox Signal* 8: 1757-1764.
- [Postgate, 1998] Postgate JR (1998). Nitrogen Fixation, 3rd Edition. Cambridge University Press, Cambridge UK.
- [Ramu et al., 2002] Ramu SK, Peng HM, Cook DR (2002) Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene rip1 in *Medicago truncatula*. *Mol Plant Microbe Interact* 15: 522-528.
- [Rees e Howard, 2000] Rees DC, Howard JB (2000) Nitrogenase: standing at the crossroads. *Curr Opin Chem Biology* 4: 559-566.
- [Rees et al., 2005] Rees DC, Tezcan FA, Haynes CA, Waton MY, Andrade S, Einsle O, Howard JB (2005) Structural basis of biological nitrogen fixation. *Phil Trans R Soc* 363: 971-984.
- [Reinhold-Hurek e Hurek, 1998] Reinhold-Hurek B, Hurek T (1998) Interactions of gramineous plants with *Azoarcus* spp. and other diazotrophs: Identification, localization and perspectives to study their function. *Crit Rev Plant Sci* 17: 49-54.
- [Reis et al., 1994] Reis VM, Olivares FL, Döbereiner J (1994) Improved methodology for isolation of *Acetobacter diazotrophicus* and confirmation of its endophytic habitat. *World Journal Microbiology Biotechnology* 10: 101-104.
- [Reis e Döbereiner, 1998] Reis VM, Döbereiner J (1998) Effect of high sugar concentration on nitrogenase activity of *Acetobacter diazotrophicus*. *Arch Microbiol* 171: 13-18.
- [Reitzer, 1996] Reitzer LJ (1996) *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. Neidhardt FC, Curtiss R, Ingraham JL, Lin ECC, Low KB (eds) *Am Soc Microbiol* p. 391-407. Washington, DC.
- [Ren e Paulsen, 2005] Ren Q, Paulsen IT (2005) Comparative analyses of fundamental differences in membrane transport capabilities in prokaryotes and eukaryotes. *PLoS Comput Biol* 3: 27.
- [Rivero et al., 2007] Rivero RM, Kojima M, Gepstein A, Sakakibara H, Mittler R, Gepstein S, Blumwald E (2007) Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc Natl Acad Sci USA* 104: 19631-19636.
- [Robson e Postgate, 1980] Robson RL, Postgate JR (1980) Oxygen and hydrogen in biological nitrogen fixation. *Annu Rev Microbiol* 34: 183-207.

- [Rodés, 2002] Rodés R (2002) Respiration and other physiological responses as oxygen protection in plant-nitrogen fixing endophyte interactions in sugarcane Em: Finan TM, O'Brian MR, Layzell DB, Vessey JK, Newton W (eds) Nitrogen fixation: Global Nitrogen Perspectives, p. 469. CABI Publishing, New York.
- [Rodrigues Neto et al., 1986] Rodrigues Neto J, Malavolta Júnior VA, Victor O (1986) Meio simples para o isolamento e cultivo de *Xanthomonas campestris* pv. citri tipo B. Summa Phytopathologica, Campinas 12: 16.
- [Rosconi et al., 2006] Rosconi F, Souza EM, Pedrosa FO, Platero RA, González C, González M, Batista S, Gill PR, Fabiano ER (2006) Iron depletion affects nitrogenase activity and expression of nifH and nifA genes in *Herbaspirillum seropedicae*. FEMS Microbiol Lett 258: 214-219.
- [Rouws et al., 2007] Rouws LFM, Simões-Araújo JL, Hemerly AS, Baldani JI (2007) Validation of a Tn5 transposon mutagenesis system for *Gluconacetobacter diazotrophicus* through characterization of a flagellar mutant. Arch Microbiol 189: 397-405.
- [Rutherford et al., 2000] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B (2000) Artemis: sequence visualization and annotation. Bioinformatics 16: 944-945.
- [Saier, 2000] Saier MHJr (2000) Vectorial metabolism and Evolution of transport systems. J. Bacteriology 182: 5029-5035.
- [Saier et al., 2006] Saier MHJr, Tran CV, Barabote RD (2006) TCDB: The transporter classification database for membrane transport protein analyses and information. Nucleic Acids Res 34: 181-186.
- [Salazar et al., 2001] Salazar JC, Zúñiga R, Raczniak G, Becker H, Söll D, Orellana O (2001) A dual-specific *Glu - tRNA^{Gln}* and *Asp - tRNA^{Asn}* amidotransferase is involved in decoding glutamine and asparagine codons in *Acidithiobacillus ferrooxidans*. FEBS Letters 500: 129-131.
- [Sali, 2006] Sali A (2006) MODELLER - A Program for protein structure modeling. Disponível em <http://salilab.org/modeller/manual/manual.html>. Acesso em 10 de maio de 2006.
- [Sambrook e Russell, 1986] Sambrook J, Russell DW (1986) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- [Sanger, 2006] Sanger Institute. Informatics Analysis Software: Alfresco, November 2003. Disponível em <http://www.sanger.ac.uk/Software/Alfresco>. Acesso em 10 de maio de 2006.

- [Santos et al., 2001] Santos R, Herouart D, Sigaud S, Touati D, Puppo A (2001) Oxidative burst in alfalfa – *Sinorhizobium meliloti* symbiotic interaction. *Mol Plant Microbe Interact* 14: 86-89.
- [Santos et al., 2005] Santos MBQ (2005) Montagem e finalização do genoma de *Gluconacetobacter diazotrophicus* PAL5. 105 p. Dissertação de Mestrado em Química Biológica. Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- [Santos et al., 2009] Santos CX, Tanaka LY, Wosniak J, Laurindo FRM (2009) Mechanisms and implications of reactive oxygen species generation during the unfolded protein response: roles of endoplasmic reticulum oxidoreductases, mitochondrial electron transport, and NADPH oxidase. *Antioxidants & Redox Signalling* 11: 2409-2427.
- [Saravanan et al., 2007] Saravanan VS, Kalaiarasan P, Madhaiyan M, Thangaraju M (2007) Solubilization of insoluble zinc compounds by *Gluconacetobacter diazotrophicus* and the detrimental action of zinc ion (Zn^{+2}) and zinc chelates on root knot nematode *Meloidogyne incognita*. *Letters in Applied Microbiology* 44: 235-241.
- [Schafer e Buettner, 2001] Schafer F, Buettner G (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radic. Biol. Med.* 30: 1191-212.
- [Schön et al., 1988a] Schön A, Kannagara CG, Gough S, Söll D (1988a). Protein biosynthesis in organelles requires misaminoacylation of tRNA. *Nature* 331: 187-190.
- [Schön et al., 1988b] Schön A, Hottinger H, Söll D (1988b). Misaminoacylation and transamidation are required for protein biosynthesis in *Lactobacillus bulgarius*. *Biochimie* 70: 391-394.
- [Schulze et al., 1998] Schulze J, Shi L, Blumenthal J, Samac DA, Gantt JS, Vance CP (1998) Inhibition of alfalfa root nodule phosphoenolpyruvate carboxylase through an antisense strategy impacts nitrogen fixation and plant growth. *Phytochemistry* 49: 341-346.
- [Schwartz, 2000] Schwartz I (2000) Microbial Genomics: From Sequence to Function. *Genomics*. 6: 493-495.
- [Serraj et al., 1994] Serraj R, Roy G, Drevon JJ (1994) Salt stress induces a decrease in the oxygen uptake of soybean nodules and in their permeability to oxygen diffusion. *Physiologia Plantarum* 91: 161-168.
- [Sevilla et al., 2001] Sevilla M, Burris RH, Gunapala N, Kennedy C (2001) Comparison of benefit to sugarcane plant growth and $^{15}N_2$ incorporation following inoculation of sterile plants with *Acetobacter diazotrophicus* wild-type and Nif- mutants strains, *Mol Plant Microbe Interact* 14: 358-366.

- [Sessitsch et al., 2005] Sessitsch A, Coenye T, Sturz AV, Vandamme P, Barka EA, Salles JF, Van Elsas JD, Faure D, Reiter B, Glick BR, Wang-Pruski G, Nowak J (2005) *Burkholderia phytofirmans* sp. nov., a novel plant-associated bacterium with plant-beneficial properties. *Int J Syst Evol Microbiol* 55: 1187-1192.
- [Simpson et al., 2000] Simpson, A.J. et al. (2000) The genome sequence of the plant pathogen *Xyllela fastidiosa*. *Nature* 406: 151-157.
- [Simpson e Burris, 1984] Simpson FB, Burris RHA (1984) Nitrogen pressure of 50 atmospheres does not prevent evolution of hydrogen by nitrogenase. *Science* 224: 1095-1096.
- [Bionotes, 2006] Sistema Bionotes. Riogene: Anotação do genoma de *Gluconacetobacter diazotrophicus*. Disponível em: <http://www.bionotes.com.br>. Acessado em 10 de maio de 2006.
- [Smil, 2000] Smil V (2000) *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. The MIT Press, Massachusetts.
- [Spiff e Odu, 1973] Spiff ED, Odu CTI (1973) Acetylene Reduction by *Beijerinckia* under Various Partial Pressures of Oxygen and Acetylene. *Journal of General Microbiology* 78: 207-209.
- [Stephan et al., 1991] Stephan MP, Oliveira M, Teixeira KRS, Martínez-Drets G, Döbereiner J (1991) Physiology and dinitrogen fixation of *Acetobacter diazotrophicus*. *FEMS Microbiol Lett* 77: 67-72.
- [Swings, 1992] Swings J (1992) The genera *Acetobacter* and *Gluconobacter*. Em: A. Balows, H. G. Trüper, M. Dworkin, W. Harder, and K.-H. Schleifer (ed.), *The prokaryotes. A handbook on the biology of bacteria: ecophysiology, isolation, identification, applications III*: 2268-2286. Springer-Verlag, New York.
- [Tapia-Hernández et al., 2000] Tapia-Hernández A, Bustillos-Cristales MR, Jimenez-Salgado T, Caballero-Mellado J, Fuentes-Ramires LE (2000) Natural endophytic occurrence of *Acetobacter diazotrophicus* in pineapple plants. *Microbial Ecology* 39: 549-555.
- [Tejera et al., 2004] Tejera NA, Ortega E, Rodés R, Lluch C (2004) Influence of carbon and nitrogen sources on growth, nitrogenase activity and carbon metabolism of *Gluconacetobacter diazotrophicus*. *Can J Microbiol* 50: 745-750.
- [Tejera et al., 2006] Tejera N, Ortega E, Rosa Rodés, Llucha C (2006) Nitrogen compounds in the apoplastic sap of sugarcane stem: Some implications in the association with endophytes. *Journal of Plant Physiology* 163: 80-85.

- [Terada, 2006] Terada LS (2006) Specificity in reactive oxidant signaling: think globally, act locally. *J Cell Biology* 174: 615-623.
- [Tigr, 2006] TIGR – The Institute for Genomic Research. RBSFinder. Disponível em <http://www.tigr.org/software/genefinding.shtml>. Acesso em 10 de maio de 2006.
- [Toh et al., 1994] Toh H, Kawamura T, Izui K (1994) Molecular evolution of phosphoenolpyruvate carboxylase. *Plant Cell Environ* 17: 31-43.
- [Tumbula et al., 2000] Tumbula DL, Becker HD, Chang WZ, Söll D (2000) Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* 407: 106-110.
- [UniProt Consortium, 2008] UniProt Consortium (2008) The universal protein resource (Uni-Prot). *Nucleic Acids Res* 36: 190-195.
- [Ureta et al., 1995] Ureta A, Alvarez B, Ramón A, Vera MA and Martínez-Drets G (1995) Identification of *Acetobacter diazotrophicus*, *Herbaspirillum seropedicae* and *Herbaspirillum rubrisubalbicans* using biochemical and genetic criteria. *Plant Soil* 172: 271-277.
- [Ureta e Nordlund, 2002] Ureta A, Nordlund S (2002) Evidence for Conformational Protection of Nitrogenase against Oxygen in *Gluconacetobacter diazotrophicus* by a Putative FeSII Protein. *J Bacteriol* 184: 5805-5809.
- [van Heerden et al., 2008] van Heerden PD, Kiddle G, Pellny TK, Mokwala PW, Jordaan A, Strauss AJ, de Beer M, Schluter U, Kunert KJ, Foyer CH (2008) Roles for the regulation of respiration and the oxygen diffusion barrier in soybean in the protection of symbiotic nitrogen fixation from chilling-induced inhibition and shoots from premature senescence. *Plant Physiol* 148: 316-327.
- [Vargas et al., 1994] Vargas C, Wu G, Davies AE, Downie JA (1994) Identification of a gene encoding a thioredoxin-like product necessary for cytochrome c biosynthesis and symbiotic nitrogen fixation in *Rhizobium leguminosarum*. *J. Bacteriol* 176: 4117-4123.
- [Waters e Emerich, 2000] Waters JK, Emerich DW (2000) Transport of metabolites to and from symbiosomes and bacteroids. Em: Triplett EW (ed) *Horizon Scientific Press, Wymondham*.
- [Wei e Layzell, 2006] Wei H, Layzell DB (2006) Adenylate-coupled ion movement: a mechanism for the control of nodule permeability to O₂ diffusion. *Plant Physiol* 141: 280-287.
- [Wealand et al., 1989] Wealand JL, Myers JA, Hirschberg R (1989) Changes in gene expression during nitrogen starvation in *Anabaena variabilis* ATCC 29413. *J Bacteriol.* 171: 1309-1313.

- [Werten et al., 2002] Werten PJ, Remigy HW, de Groot BL, Fotiadis D, Philippsen A, Stahlberg H, Grubmuller H, Engel A (2002) Progress in the analysis of membrane protein structure and function. *FEBS Lett* 529: 65-72.
- [White e Bayley, 1972] White BN, Bayley ST (1972) Further codon assignments in an extremely halophilic bacterium using a cell-free protein-synthesizing system and a ribosomal binding assay. *Canadian Journal of Biochemistry* 50: 600-609.
- [Wilcox, 1969] Wilcox M (1969) γ -phosphoryl ester of *Glu-tRNA^{Gln}* as an intermediate in *Bacillus subtilis* glutaminyl-tRNA synthesis. *Cold Spring Harbor Symp. Quant. Biol.* 34: 521-528.
- [Wilcox e Nirenberg, 1968] Wilcox M, Nirenberg M (1968) Transfer RNA as a cofactor coupling amino acid synthesis with that of protein. *Proceedings of the National Academy of Sciences of the USA* 61: 229- 236.
- [Wolfson et al., 1998] Wolfson AD, Pleiss JA, Uhlenbeck OC (1998) A new assay for tRNA aminoacylation kinetics. *RNA* 4: 1019-1023.
- [Wu-blast, 2006] WU-BLAST – Washington University BLAST Archives-BLAST Manual. Disponível em <http://blast.wustl.edu/doc/blast1.pdf>. Acesso em 10 de maio de 2006.
- [Wu et al., 2002] Wu CH, Huang H, Arminski L, Castro-Alvear J, Chen Y, Hu ZZ, Ledley RS, Lewis KC, Mewes HW, Orcutt BC, Suzek BE, Tsugita A, Vinayaka CR, Yeh LS, Zhang J, Barker WC (2002) The protein information resource: an integrated public resource of functional annotation of proteins. *Nucleic Acids Res* 30: 35-37.
- [Yamada et al., 1997] Yamada Y, Hoshino K, Ishikawa T (1997) The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA: The elevation of the subgenus *Gluconoacetobacter* to the generic level. *Biosci Biotechnol Biochem* 61: 1244-1251.
- [Yamada et al., 1998] Yamada Y, Hoshino K, Ishikawa T (1998) *Gluconoacetobacter* nom. corrig. (*Gluconoacetobacter* [sci]). Em: Validation of publication of new names and new combinations previously effectively published outside the IJSB, List no. 64. *Int J Syst Bacteriol* 48: 327-328.
- [Young, 1992] Young JPW (1992) Phylogenetic classification of nitrogen-fixing organism. Em: Stacey G, Burris RH, Evans HJ (eds) *Biological Nitrogen Fixation*, p. 43-86. Chapman & Hall, New York.
- [Zanin et al., 2000] Zanin GM, Santana CC, Bon EP, Giordano RC, de Moraes FF, Andrietta SR, de Carvalho Neto CC, Macedo IC, Fo DL, Ramos LP, Fontana JD (2000) Brazilian bioethanol program. *Appl Biochem Biotechnol* 84: 1147-1161.

Apêndice

- Manuscrito 1 - Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5
- Manuscrito 2 - Antioxidant pathways are up-regulated during biological nitrogen fixation to prevent ROS-induced nitrogenase inhibition in *Gluconacetobacter diazotrophicus*
- Sylvia Maria Campbell Alquéres - *Curriculum vitae*

Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5

Marcelo Bertalan¹, Rodolpho Albano², Vânia de Pádua³, Luc Rouws⁴, Cristian Rojas¹, Adriana Hemerly^{1,12}, Kátia Teixeira⁴, Stefan Schwab⁴, Jean Araujo⁴, André Oliveira⁴, Leonardo França¹, Viviane Magalhães¹, Sylvia Alquéres¹, Alexander Cardoso¹, Wellington Almeida¹, Marcio Martins Loureiro¹, Eduardo Nogueira^{3,11}, Daniela Cidade², Denise Oliveira², Tatiana Simão², Jacyara Macedo², Ana Valadão², Marcela Dreschsel⁴, Flávia Freitas², Marcia Vidal⁴, Helma Guedes⁴, Elisete Rodrigues⁴, Carlos Meneses⁴, Paulo Brioso⁵, Luciana Pozzer⁵, Daniel Figueiredo⁵, Helena Montano⁵, Jadier Junior⁵, Gonçalo de Souza Filho⁶, Victor Martin Quintana Flores⁶, Beatriz Ferreira⁶, Alan Branco⁶, Paula Gonzalez⁷, Heloisa Guillobel⁷, Melissa Lemos⁸, Luiz Seibel⁸, José Macedo⁸, Marcio Alves-Ferreira⁹, Gilberto Sachetto-Martins⁹, Ana Coelho⁹, Eidy Santos⁹, Gilda Amaral⁹, Anna Neves⁹, Ana Beatriz Pacheco¹⁰, Daniela Carvalho¹⁰, Letícia Lery¹⁰, Paulo Bisch¹⁰, Shaila C Rössle¹⁰, Turán Ürményi¹⁰, Alessandra Rael Pereira², Rosane Silva¹⁰, Edson Rondinelli¹⁰, Wanda von Krüger¹⁰, Orlando Martins¹, José Ivo Baldani⁴ and Paulo CG Ferreira*^{1,12}

Address: ¹Instituto de Bioquímica Médica, UFRJ, CCS, Bloco D, subsolo 21491-590 Rio de Janeiro, Brazil, ²Departamento de Bioquímica, Instituto de Biologia Roberto Alcântara Gomes, UERJ, Blv 28 de Setembro, 87, fundos, 4 andar, Vila Isabel, Rio de Janeiro, RJ 20551-013, Brazil, ³Laboratório de Tecnologia em Bioquímica e Microscopia, Centro Universitário Estadual da Zona Oeste, Rio de Janeiro, 23070-200, Brazil, ⁴Embrapa Agrobiologia BR465, Km 07 Seropédica Rio de Janeiro, 23851-970, Brazil, ⁵Instituto de Biologia, Departamento de Entomologia e Fitopatologia, Universidade Federal Rural do Rio de Janeiro Cx Postal 74585/BR 465, KM 07, Seropédica, RJ 23851-970, Brazil, ⁶Lab. Biotecnologia- Centro de Biociências e Biotecnologia Universidade Estadual do Norte Fluminense- Av. Alberto Lamego 2000 Campos dos Goytacazes RJ 28013-620, Brazil, ⁷Departamento de Biofísica e Biometria, Instituto de Biologia Roberto Alcântara Gomes UERJ, Blv 28 de Setembro, 87, fundos, 4 andar, Vila Isabel, Rio de Janeiro, RJ 20551-013, Brazil, ⁸Departamento de Informática - Pontifícia Universidade Católica do Rio de Janeiro Rua Marquês de S. Vicente, 225, Rio de Janeiro 22453-900, Brazil, ⁹Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Cx Postal 68011, Rio de Janeiro, RJ 21941-617, Brazil, ¹⁰Instituto de Biofísica Carlos Chagas Filho Universidade Federal do Rio de Janeiro, CCS, Cidade Universitária, Rio de Janeiro, RJ21.949-900, Brazil, ¹¹Laboratório de Biologia Molecular, Departamento de Genética e Biologia Molecular, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, RJ 22290-240, Brazil and ¹²Laboratório de Biologia Molecular de Plantas, Instituto de Pesquisas do Jardim, Botânico do Rio de Janeiro, 22460-030 Rio de Janeiro, RJ, Brazil

Email: Marcelo Bertalan - bertalan@cbs.dtu.dk; Rodolpho Albano - albano@uerj.br; Vânia de Pádua - vaniadepadua@gmail.com; Luc Rouws - lucrouws@hotmail.com; Cristian Rojas - rojas@bioqmed.ufrj.br; Adriana Hemerly - hemerly@bioqmed.ufrj.br; Kátia Teixeira - katia@cnpab.embrapa.br; Stefan Schwab - stephan@cnpab.embrapa.br; Jean Araujo - jean@cnpab.embrapa.br; André Oliveira - almoliva@yahoo.com.br; Leonardo França - leomegaf@bioqmed.ufrj.br; Viviane Magalhães - vmagalhaes@bioqmed.ufrj.br; Sylvia Alquéres - salqueres@gmail.com; Alexander Cardoso - amcardoso@bioqmed.ufrj.br; Wellington Almeida - walmeida@inpi.gov.br; Marcio Martins Loureiro - loureiro@bioqmed.ufrj.br; Eduardo Nogueira - eduardonogu@gmail.com; Daniela Cidade - daniela_cidade_6@hotmail.com; Denise Oliveira - dneves_oliveira@yahoo.com; Tatiana Simão - tsimao@terra.com; Jacyara Macedo - jacyara@uerj.br; Ana Valadão - valadao@gmail.com; Marcela Dreschsel - marceladreschsel@yahoo.com.br; Flávia Freitas - flaviaalvim_dutra@yahoo.com.br; Marcia Vidal - marcia@cnpab.embrapa.br; Helma Guedes - helmavg@gmail.com; Elisete Rodrigues - bioliza@yahoo.com; Carlos Meneses - chmeneses@gmail.com; Paulo Brioso - paulobri@ufrj.br; Luciana Pozzer - luciana.pozzer@agricultura.gov.br; Daniel Figueiredo - danvazfig@ig.com.br; Helena Montano - hmontano@ufrj.br; Jadier Junior - jadier_junior@yahoo.com; Gonçalo de Souza Filho - goncalos@uenf.br; Victor Martin Quintana Flores - martin@uenf.br; Beatriz Ferreira - ferreirabs@uenf.br; Alan Branco - alantb@uenf.br; Paula Gonzalez - paulinhagonzalez@yahoo.com.br; Heloisa Guillobel - hguillo@uerj.br; Melissa Lemos - melissa@inf.puc-rio.br; Luiz Seibel - seibel@inf.puc-rio.br; José Macedo - jmacedo@inf.puc-rio.br; Marcio Alves-Ferreira - alvesferreira@pesquisador.cnpq.br; Gilberto Sachetto-

Martins - sachetto@biologia.ufrj.br; Ana Coelho - coelho@biologia.ufrj.br; Eidy Santos - eidyos@aol.com; Gilda Amaral - gildabio@biologia.ufrj.br; Anna Neves - annacrisnb@yahoo.com.br; Ana Beatriz Pacheco - biafp@biof.ufrj.br; Daniela Carvalho - dpcarval@yahoo.com; Letícia Lery - lery@biof.ufrj.br; Paulo Bisch - pmbisch@biof.ufrj.br; Shaila C Rössle - shaila@biof.ufrj.br; Turán Úrményi - turmenyi@biof.ufrj.br; Alessandra Rael Pereira - alessandra@gmail.com; Rosane Silva - silvaros@biof.ufrj.br; Edson Rondinelli - edrondin@biof.ufrj.br; Wanda von Krüger - wmvk@biof.ufrj.br; Orlando Martins - omartins@bioqmed.ufrj.br; José Ivo Baldani - ibaldani@cnpab.embrapa.br; Paulo CG Ferreira* - paulof@bioqmed.ufrj.br
* Corresponding author

Published: 23 September 2009

Received: 13 January 2009

BMC Genomics 2009, 10:450 doi:10.1186/1471-2164-10-450

Accepted: 23 September 2009

This article is available from: <http://www.biomedcentral.com/1471-2164/10/450>

© 2009 Bertalan et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: *Gluconacetobacter diazotrophicus* Pal5 is an endophytic diazotrophic bacterium that lives in association with sugarcane plants. It has important biotechnological features such as nitrogen fixation, plant growth promotion, sugar metabolism pathways, secretion of organic acids, synthesis of auxin and the occurrence of bacteriocins.

Results: *Gluconacetobacter diazotrophicus* Pal5 is the third diazotrophic endophytic bacterium to be completely sequenced. Its genome is composed of a 3.9 Mb chromosome and 2 plasmids of 16.6 and 38.8 kb, respectively. We annotated 3,938 coding sequences which reveal several characteristics related to the endophytic lifestyle such as nitrogen fixation, plant growth promotion, sugar metabolism, transport systems, synthesis of auxin and the occurrence of bacteriocins. Genomic analysis identified a core component of 894 genes shared with phylogenetically related bacteria. Gene clusters for *gum*-like polysaccharide biosynthesis, *tad* pilus, quorum sensing, for modulation of plant growth by indole acetic acid and mechanisms involved in tolerance to acidic conditions were identified and may be related to the sugarcane endophytic and plant-growth promoting traits of *G. diazotrophicus*. An accessory component of at least 851 genes distributed in genome islands was identified, and was most likely acquired by horizontal gene transfer. This portion of the genome has likely contributed to adaptation to the plant habitat.

Conclusion: The genome data offer an important resource of information that can be used to manipulate plant/bacterium interactions with the aim of improving sugarcane crop production and other biotechnological applications.

Background

In recent years, concerns about fossil fuel supplies and prices have motivated the search for renewable biofuels. With the existing technologies and current costs of fuel transportation, ethanol from sugarcane is the most viable alternative. In some countries, including Brazil, sugarcane is planted with low amounts of nitrogen fertilizers and there is evidence that the use of low levels of nitrogen can be compensated by Biological Nitrogen Fixation (BNF) [1]. Although several organisms are capable of contributing to BNF, it has been shown that the diazotroph Alphaproteobacteria *Gluconacetobacter diazotrophicus* Pal5 (GDI), present in large numbers in the intercellular space

of sugarcane roots, stem and leaves, fixes N₂ inside sugarcane plants, without causing apparent disease [2,3]. Remarkable characteristics of this bacterium are the acid tolerance, the inability to use nitrate as sole nitrogen source and the ability to fix nitrogen in the presence of ammonium in medium with high sugar concentration [2]. Although isolation of GDI from the sugarcane rhizosphere has been reported [4], its poor survival soil and complete absence in soil samples collected between sugarcane rows strongly support the endophytic nature of this nitrogen fixing bacterium [5-7]. In addition to BNF, GDI has other characteristics that contribute to its biotechnological importance: 1-) A *nif*- mutant enhances plant

growth, particularly in roots, indicating that GDI secretes plant growth-promoting substances [8]; 2-) It produces a lysozyme-like bacteriocin that inhibits the growth of the sugarcane pathogen *Xanthomonas albilineans* [9]; 3-) It has antifungal activity against *Fusarium sp.* and *Helminthosporium carbonum* [10]; 4-) GDI promotes an increase in the solubility of phosphate and zinc [11]. Besides its biotechnological features, the genome is especially interesting because it is the third diazotrophic endophytic bacteria to be completely sequenced. The first two diazotrophic endophytes to be sequenced, *Azoarcus sp.* strain BH72 [12] and *Klebsiella pneumoniae* 342 [13], belong to the Betaproteobacteria and Gammaproteobacteria classes, respectively. Thus, the genome of GDI is the first to be completely sequenced from Alphaproteobacteria class. Here we report the complete genome sequence of the *G. diazotrophicus* strain Pal5. Sequence analysis shows the existence of a large accessory genome, probably originated by extensive Horizontal Gene Transfer (HGT). Moreover, experimental results reveal differences in Genomic Islands (GI) among *G. diazotrophicus* strains. The knowledge of the metabolic routes, organization and regulation of genes involved in nitrogen fixation, establishment of successful plant association and other processes should allow a better understanding of the role played by this bacterium in plant-bacteria interaction.

Results

Overview of the *G. diazotrophicus* PAL5 genome

The complete genome of GDI is composed of one circular chromosome of 3,944,163 base pairs (bp) with an average G+C content of 66.19%, and two plasmids of 38,818 and 16,610 bp, respectively (table 1). The circular chromosome has a total of 3,864 putative coding sequences (CDS), with an overall coding capacity of 90.67%. Among

Table 1: General features of the *G. diazotrophicus* PAL5 genome.

Features	
Size, bp	3,999,591
G+C content, %	66%
Coding sequences	3,938
Functional assigned	2,861
Insertion Sequences (IS)	223
Pseudo genes	83
Conserved and hypothetical proteins	1,077
% of the genome coding	90
Average length, bp	947
%ATG initiation codons	2,809
%GTG initiation codons	681
%TTG initiation codons	440
RNA elements	
rRNA	4 × (16S-23S-5S)
tRNAs	55

the predicted genes, 2,861 were assigned a putative function, and 1,077 encode hypothetical proteins. Regarding noncoding RNA genes, 12 rRNAs (four rRNA operons) and 55 tRNAs were identified. The larger plasmid (pGD01) has 53 CDS; approximately 70% encode hypothetical or conserved hypothetical proteins and five encode proteins involved in plasmid-related functions. The remaining 11 CDS encode putative components of the Type IV secretion system (T4SS). The small plasmid (pGD02) has 21 CDS, and around 50% are hypothetical proteins.

Although today the genome databases have more than 800 complete microbial genomes, only nine are endophytes (*Azoarcus sp.* BH72, *Burkholderia phytofirmans* PsJN, *Enterobacter sp.* 638, *Methylobacterium populi* BJ001, *Pseudomonas putida* W619, *Serratia proteamaculans* 568, *Klebsiella pneumoniae* 342, *Stenotrophomonas maltophilia* R551-3 and *Gluconacetobacter diazotrophicus* Pal5) [14]. The complete genomes of endophytic bacteria reveal remarkably few mobile elements in its genome (Additional file 1), an observation that led to the proposal that this could denote an adaptation to a more stable life style [12]. In contrast, GDI contains 190 transposable elements, more than any other endophytic bacteria (Additional file 1). The large number of mobile elements could be a signature of a recent evolutionary bottleneck and consequent relaxation of selection, perhaps due to a recent change in niche [15]. Alternatively, because GDI is found in low frequency at the rhizosphere, the transposable elements could have been acquired from other bacteria inhabiting the same environment. In order to identify possible specific characteristics of the genome, the Predicted Highly Express Genes (PHX) genes were identified [16]. PHX analysis identified 658 CDS (17% of the chromosome) in GDI with E(g) (general expression level) > 1.0. Combining this information with the proteomic results [17], which sequenced peptides from 541 genes, we identified 318 of these genes PHX. As expected, ribosomal proteins, translation/transcription factors and chaperone/degradation genes are among the top 30 E(g) values within the 318 CDS, (Additional file 2). However, some unexpected CDS also appear as PHX. For instance, there are 50 transporter proteins or transporter-related proteins with high E(g) value, of which 27 are putative ABC transporter proteins and six are putative TonB-dependent receptors. The genome has two ammonium transporter proteins (GDI0706 and GDI2352) and both with high E(g) values. Two other proteins related to ammonium metabolism are also PHX: a putative glutamate-ammonia-ligase adenylyltransferase (GDI3425) and a putative histidine-ammonia-lyase (GDI0550). This finding is consistent with the fact that ammonium is the preferred nitrogen source for GDI when it is available.

Core and accessory regions

Analysis of the core and accessory regions of GDI is important in order to understand its evolution and adaptation to the plant environment [18]. Even though Pal5 is the first *Gluconacetobacter diazotrophicus* strain to be sequenced, it is possible to identify the core genome by comparing with closely related species. The closest completed genomes available in the database were identified by phylogenetic analysis (Additional file 3). These include *Acidiphilium cryptum* JF-5 (ACC), *Gluconobacter oxydans* 621H (GOX) and *Granulibacter bethesdensis* CGDNIH (GRB). Using quartops analysis (quartets of orthologous proteins [19]) we identified 894 CDS as core. Most of these CDS are related to metabolism, information transfer and energy metabolism, as illustrated in figure 1. As CDS with low GC3 (G+C content of synonymous third position) are potential accessory genes, the mean and standard deviation of the non-quartops were used as cut-offs to identify possible accessory genes. We found that 1,352 CDS have a GC3 percentage lower than 80% (figure 2). Interpolated Variable Order Motifs [20] (IVOMs) were used to complement the accessory genome analysis,

revealing that 1,164 CDS have an "Alien score" greater than the threshold, 11,134. The group of CDS in common between GC3 and IVOMs (851 CDS) was used to define the accessory genes in the genome. The percentages of conserved hypothetical proteins, hypothetical proteins, phage/IS elements and pseudo genes are higher in the putative accessory regions than in the core regions and in the genome (figure 1), suggesting that the putative accessory regions have been transferred horizontally into the genome. Overall, the putative accessory regions cover approximately 24% of the GDI genome and are separated into 28 distinct regions, of which seven are classified as phage regions (Additional file 4). A third and completely independent method, PHX, also supports the assignment of the predicted accessory regions (figure 3).

Genome Islands: Variation among G. diazotrophicus strains

Because HGT is an important source of intra-specific genetic variation in bacteria [21], we investigated whether there are differences in putative genome islands among 19 *G. diazotrophicus* strains and one *G. johannae* strain, using

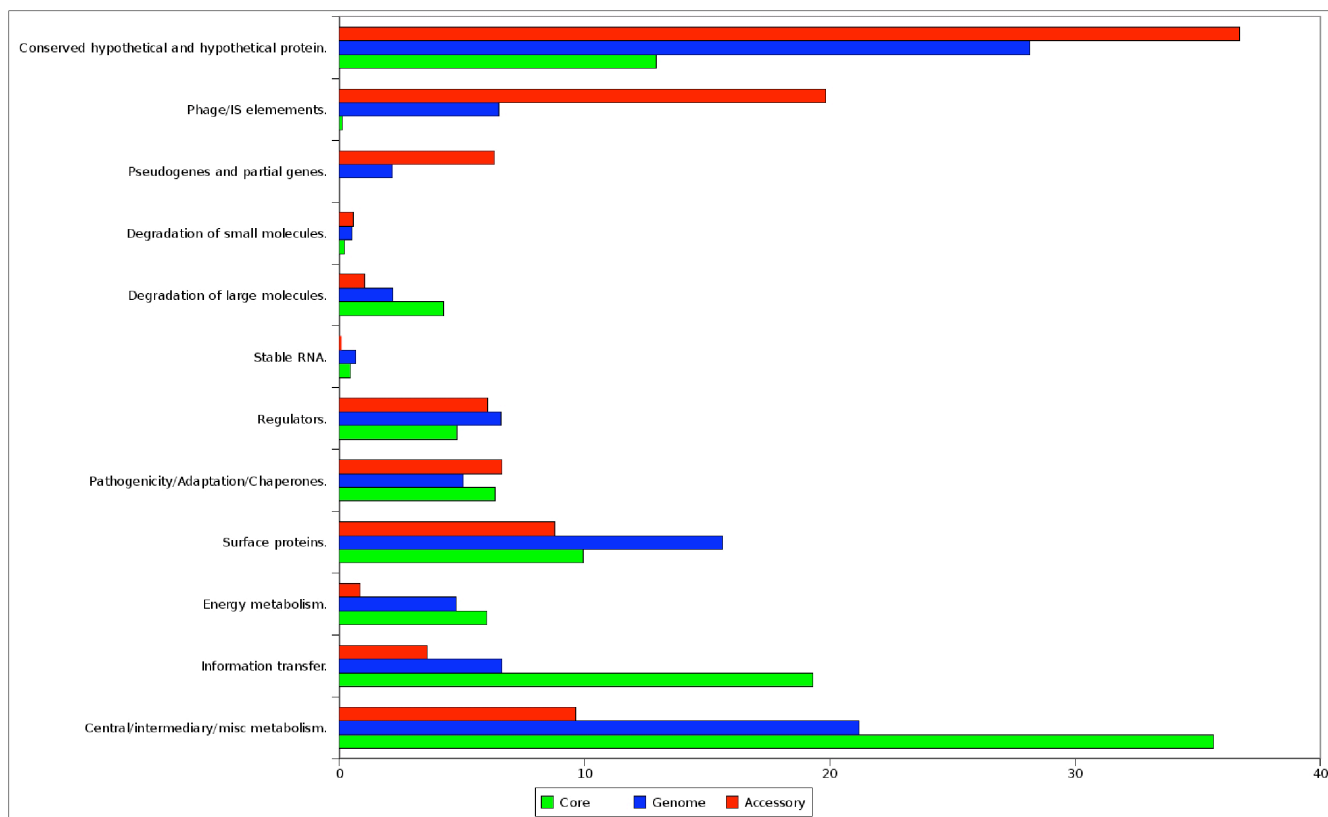


Figure 1
Distribution of gene class by groups. Percentage of gene class in three groups: Whole genome (blue), core regions (green) and accessories regions (red). The group energy metabolism includes glycolysis, electron transport. Information transfer includes transcription, translation and DNA/RNA modification. Surface class includes inner and outer membrane, secreted proteins, and lipopolysaccharides.

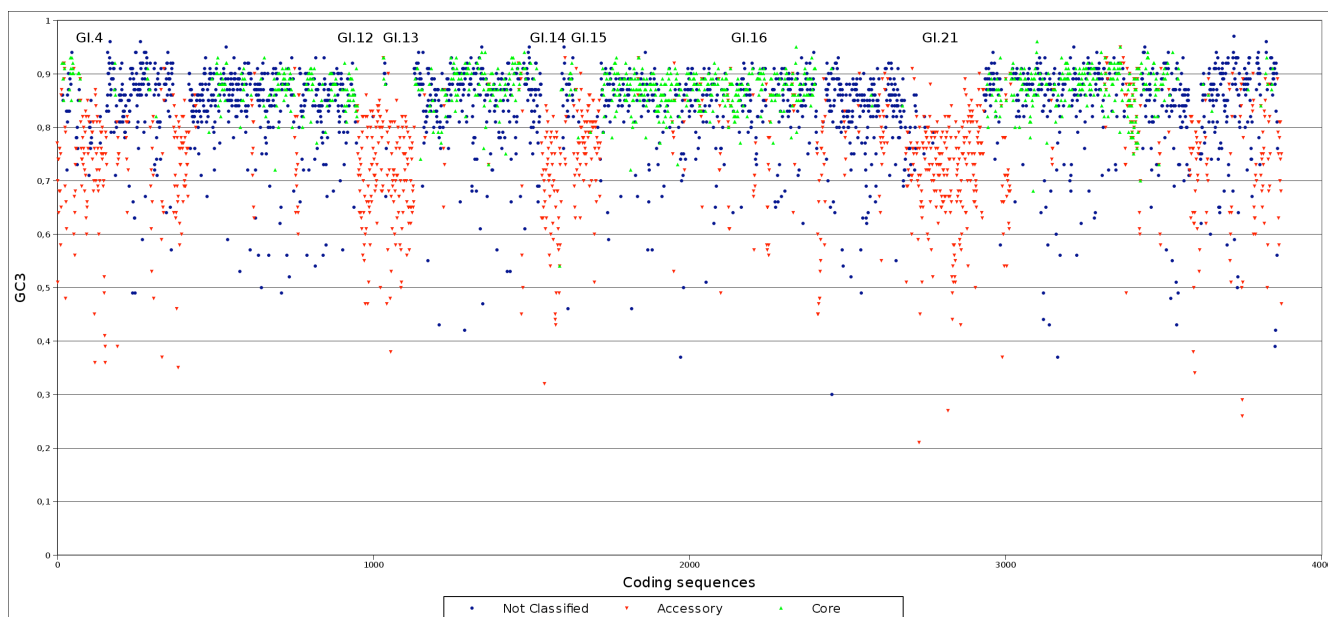


Figure 2

GC3 analysis of all genes in the chromosome. Each spot represents a gene in the chromosome. In red are the genes that were classified as accessories by the IVOM method. In green are the genes classified as core by quartops analysis. In blue are the genes that were not classified as core or accessories.

PCR with primers designed against 39 single-copy genes in 20 Genome Islands (GIs), and 17 CDS from the core genome. There was a complex variation among the strains, with gene content of eleven GIs - 1, 3, 7, 8, 9, 11, 15, 16, 17, 18, 19 - either almost entirely conserved or less than 50% variable (Additional file 5). In two GIs - 12 and 14 - there was high variability in a group of genes, while other genes were conserved in most strains. The remaining seven GIs, representing approximately 7% of the genome, were highly variable, especially GIs 4 and 21, which are 78 and 242 kb long, and encode 80 and 242 CDS, respectively. Furthermore, a considerable number of CDS in these two GIs encode genes involved in processes that could confer a competitive edge, such as oxidative stress, proteases, biosynthesis of antimicrobial agents, amino acid metabolism and secondary metabolites, as well as a large number of transport systems and transcriptional regulators. Both GI4 and GI21 also contain complete copies of the T4SS operon. As it has been suggested that T4SS can increase host adaptability in *Bartonella* [22], we suspected that they could be a source of intraspecific variation among *G. diazotrophicus* strains. A Southern blot used to probe the *trbE* gene shows that indeed the T4SS copy number varies from one to four depending on the strain (Additional file 6). These GIs could be especially important for bacterial adaptation to the endophytic lifestyle and may confer adaptation advantages to *G. diazotrophicus*

in comparison with other microbes that colonize the same niche.

General Comparison

As the experimental results support the prediction of accessory regions in GDI, another interesting question concerns which regions of the genome resembles genomes from the database. For this purpose, a Reciprocal Best Hits (RBH) comparison was used [23]. The RBH analysis indicates that only 2,966 CDSs of GDI generate a hit with a completed bacterial genome. Among them, 2,470 CDSs have best hit with the Alphaproteobacteria class, 190 with the Betaproteobacteria class, 188 CDS with the Gammaproteobacteria class and 118 with other groups. The distribution of all RBHs demonstrated that even genes from phylogenetically distant related organisms can exhibit high percent identity (Additional file 7). The organism with the highest number of best hits is GOX, with 1,099. However, in figure 1, it is possible to observe that most of the hits occur in core regions. In the three organisms closest to GDI, around 90% of the best hits occur in core regions, with 10% in accessories regions. On the other hand, among rhizobiales and other Alphaproteobacteria orders, 56% of the best hits occur in core regions and 44% in accessory regions (Additional file 8). Curiously, complete genomes from the Betaproteobacteria class, Gammaproteobacteria class and other groups

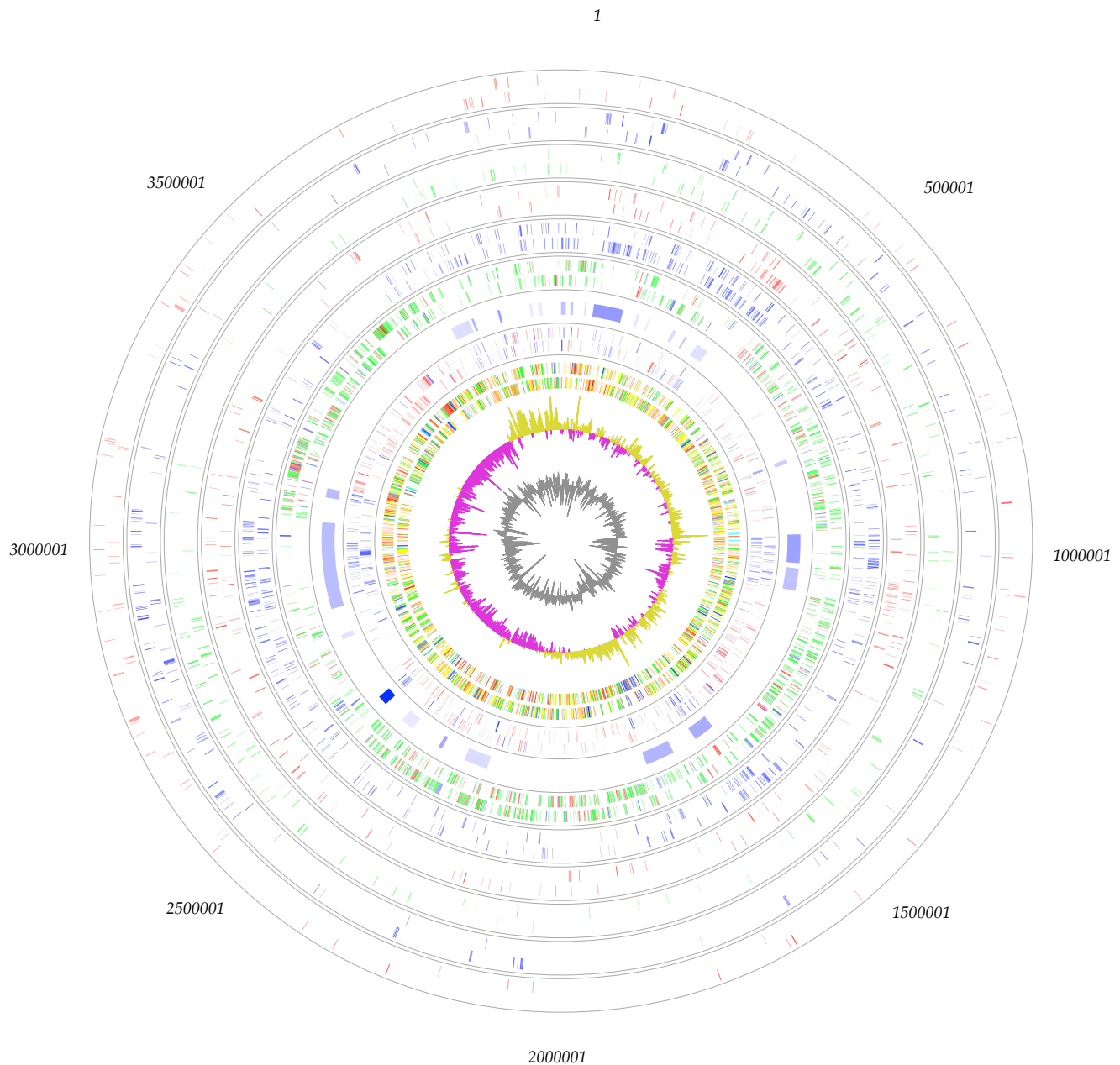


Figure 3

Circular representation of *G. diazotrophicus* PAL5 chromosome. From inside to outside. 1-) GC Content. 2-) GC Skew. 3-) Annotation, colors defined by class, see Methods. 4-) Predicted Highly Expressed genes; in blue genes classified as "Alien" and in red genes classified as putative highly expressed. 5-) Accessory regions determined by GC3 and IVOM. 6-) Reciprocal best hits results, in green from *G. oxydans* 621H, in blue genes from *A. cryptum* JF-5 and in red genes from *G. bethesdensis* CGDNIH. 7-) Reciprocal Best Hits (RBH) with all complete genomes from the order rhizobiales. 8-) RBH with all other complete genomes from Alphaproteobacteria class; 9-) RBH with all complete genomes from Betaproteobacteria class. 10-) RBH with all complete genomes from Gammaproteobacteria class. 11-) RBH with all other complete genomes.

have a significant number (65-70%) of RBHs in core regions, and low percentage (30-35%) in accessory regions. In addition, the number of RBHs with phytopathogenic organisms is higher in Betaproteobacteria and Gammaproteobacteria than in Alphaproteobacteria (68%, 55% and 8%, respectively).

Comparisons with other endophytic bacteria

Currently, there are only nine complete genome sequences of endophytic bacteria, and all are Proteobacteria. Using the complete genomes, we searched for common and exclusive CDS among endophytic bacteria in order to identify genes that could explain the endophytic capacity. However, we found only five CDS that are exclusively conserved (Additional file 9). The comparison among the endophytic organisms indicates that GDI has more CDS exclusively conserved with *Methylobacterium populi* BJ001 (133 genes) than with the others, which is consistent with the fact that *M. populi* BJ001 is also an Alphaproteobacteria. Most of these genes (Additional file 10) occur in an accessory region (GI4, GI9, GI12, GDI13, GDI14, GDI19 and GI21), and many are putative transcriptional regulators and putative T4SS (Additional file 9), which could also be involved in bacteria-host interactions. We also searched for exclusively conserved CDS between GDI and *Azoarcus* sp. BH72, as these two bacteria are currently the only diazotrophs among the endophytes sequenced. The result confirmed the presence in both endophytes of the *nif* cluster (figure 2, around 0.5 MB) and genes from the putative *gum* cluster are only conserved within *Azoarcus* sp. BH72 and GDI (Additional file 10). An assessment of the classes and frequency of signaling CDS in both diazotrophs shows that *Azoarcus* sp. BH72 has acquired a far more complex set of regulators (Additional file 11). In contrast, GDI has many more transport systems than *Azoarcus* sp. BH72 (Additional file 12). Altogether, the strategy developed by GDI to interact with plants seems to be more similar to *Methylobacterium populi* BJ001 than to other endophytes. However, the result suggests that there is not only one strategy and probably there are different ways in which bacteria can interact with plants.

After we completed this work, a second genome sequence of *Gluconacetobacter diazotrophicus* strain Pal5 was deposited. We carried out extensive comparisons between the two sequences. The comparison is summarized in Additional file 13. The results show significant differences between the two versions. GDI-BR has 309 more CDS than GDI-US, although this number is significantly reduced when small ORFs are annotated as CDSs in GDI-US. Likewise, the number of unique genes in both genomes decreases from 747 and 438 to 624 and 110, respectively, when the small CDSs are taken into account. The results show that the transposases, integrases and

hypothetical proteins can explain the majority of the differences between the two sequences. Furthermore, 67% of the genes unique to GDI-BR are located in Genome Islands. On the other hand, 85% BBH among the two sequences are found outside the GIs. The results of the genomic comparisons between the two sequences are compatible with the PCR results reported here, that showed that most of the genic differences among GDI strains are situated in the GIs. Furthermore, when GIs from the two sequences are compared, most of the genic variation is found in the same more variable GIs (data not shown). Altogether, these analyzes suggest that the two sequences deposited as *G. diazotrophicus* Pal5 strain may represent either two different strains or a fast diverging strain.

In addition, our results were corroborated by at least three independent approaches. First, Southern Blot analyzes confirmed that the genomic sequence we have deposited has 4 copies of the TSS4 secretion system. Furthermore, PCR with primers that amplified genes in the GIs verified the presence of all CDS in our sequence, while some like GDI2782 which encodes a putative H(+)/Cl(-) exchange transporter, is absent from the second sequence. Finally, over 500 CDS in our sequence were validated by proteomics [17]. Some of these CDS may confer unique biological properties and competitiveness to *Gluconacetobacter diazotrophicus* Pal5, such as a Bacteriocin (GDI0415). Additional file 14 contains the list of Blast Best Hits among the two *Gluconacetobacter diazotrophicus* Pal5 genomic sequences, a list of unique CDS found in chromosome from GeneBank file CP001189 and a list of unique genes found in chromosome from GeneBank file AM889285 (this work).

Genome Features in Core Regions

Osmotolerance

GDI supports high sugar concentrations, being able to tolerate up to 30% sucrose, but is sensitive to salt [24]. This shows its adaptation to sugarcane tissues, where the sucrose content is frequently high. Several osmoprotection systems were found (figure 4). There is a Kdp sensor system *kdpABCDE*, which regulates potassium flux (GDI1564-1568) [25]. One putative proline/betaine transporter gene was detected (GDI2530), but transporter genes *proU*, *betT* and *opuA* were not found. Pathways for glycine/betaine production are incomplete and genes necessary for conversion from choline to betaine are absent. The GDI genome harbors three Dpp ABC transporters that facilitate the uptake of di- and tripeptides (GDI0246-0250, GDI0454-0458 and GDI3540-3544). Two ORFs encoding a DtpT transporter, also involved in the uptake of di- and tripeptides, are present (GDI3819 and GDI0829). The presence of *otsA*, *otsB* and *treA* homologs (GDI0917, GDI0916 and GDI1341) suggests that GDI

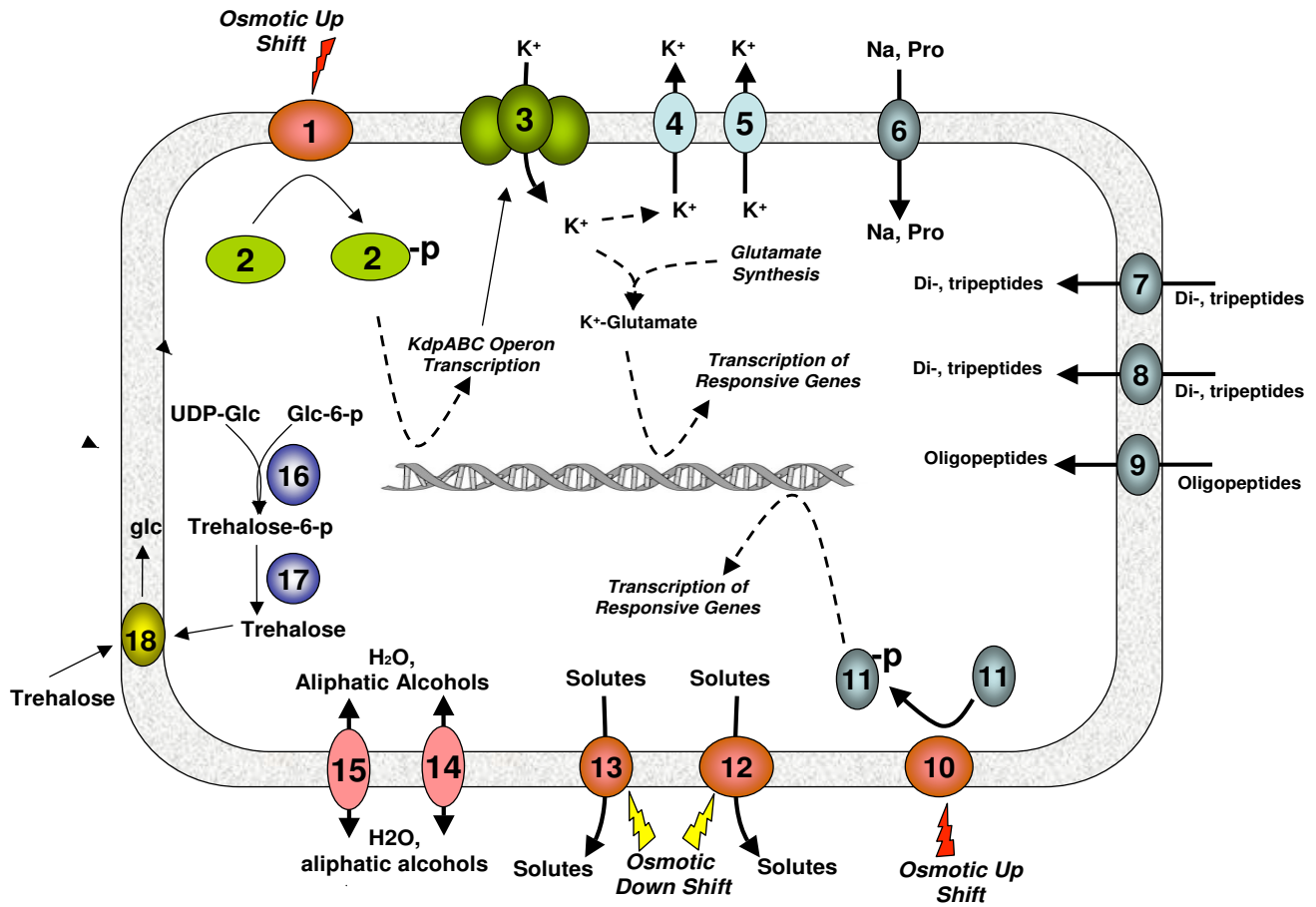


Figure 4
Osmotolerance mechanisms in *G. diazotrophicus*. Osmotolerance mechanisms in *G. diazotrophicus*. (1) Sensor protein *kdpD* (GDI1564). (2) Transcriptional regulatory protein *kdpE* (GDI1565). (3) Potassium ABC transporter (*kdpABC* transporter; GDI1566-1568). (4) Glutathione-regulated system protein *kefB* (GDI0899) and (5) *kefC* (GDI2585). (6) Proline/betaine transporter (GDI2530). (7) Dpp ABC transporters for di- and tripeptides (GDI0246-GDI0250, GDI0454-GDI0458 and GDI3540-GDI3544). (8) Transporter *dtpT*, (GDI3819 and GDI0829). (9) Oligopeptide transporter (*Opt*; GDI3108). (10) Sensor kinase *EnVZ* (GDI3087). (11) *OmpR* (GDI3088). (12) Large Conductance MS channel *mscL* (GDI1732). (13) Small conductance MS channel, *mscS*, (GDI0793, GDI1149, GDI1789, and GDI3802). (14) *glpRDFK* (GDI1751-1754). (15) *glpDKF* (GDI0262, GDI0266, and GDI0267). (16) *otsA* GDI0917. (17) *otsB* GDI0916. (18) Periplasmic trehalase (*treA* GDI1341). The function of the proteins was verified by BLAST and motif searches of the corresponding CDS against public databases.

may synthesize and use the osmolytic disaccharide trehalose, although experiments on solid culture medium have shown that GDI is able to grow poorly on trehalose as a carbon source (data not shown). The hyperosmotic sensing in GDI may occur through the two-component system *envZ/ompR* (GDI3087 and GDI3088). However, the *envZ*-regulated porins *ompF* and *ompC* are not present. In bacteria, two porins (*aqpZ* and *glpF*) regulate the movement of water and aliphatic alcohols across cell membranes [26]. Homologs of *aqpZ* are missing in GDI, although two sets of glyceroporin genes were found in two clusters: one containing *glpRDFK* (GDI1751-1754) and the other composed of *glpDKF* (GDI0262, GDI0266, and GDI0267).

The mechanisms shown in figure 4 and discussed here are similar to those found in bacteria without the high level of tolerance to high sugar concentrations observed for *G. diazotrophicus*. Therefore, unknown mechanisms that protect the bacteria specifically against high sugar concentrations may act in GDI. However, GDI seems to have a larger number of isoforms of enzymatic systems involved in osmotolerance. These differences may be explained by the different niches inhabited by GDI and *Azoarcus sp* BH72. While GDI is found in plants with elevated concentration of sugars, *Azoarcus sp* BH72 lives in association with plants that do not accumulate carbon sources in high concentration in vegetative tissues, rice and Kallar grass,

and thus *Azoarcus sp* BH72 may not need a large number of enzymes.

Acid tolerance

GDI has high tolerance to low pH and organic acids and is able to fix nitrogen at pH values as low as 2.5 [27]. The acidophile *Acetobacter aceti* has an unusual citric acid cycle (CAC) that is important for acetic acid resistance at low pH [28]. Genome analyses revealed the presence in the GDI genome of homologs of the alternative *A. aceti* citrate synthase gene *aarA* (GDI1830) and the gene for an acetyl-CoA hydrolase family protein gene with succinyl-CoA:acetate CoA-transferase activity, called *aarC* (GDI1836). In GDI, the *aarAC* homologs occur in a cluster similar to that of *A. aceti*, contrasting with the organization of these genes in non-acidophilic species, thus indicating that the same mechanisms of acid tolerance involving the CAC may be acting in both organisms. We also found a homolog of an ABC-transporter gene *aatA* (GDI1739) that, in *A. aceti*, constitutes an organic acid efflux pump mediating resistance to several acids [29]. An unusual observation is the presence in the GDI genome of two copies of the chaperonin genes *groES* (GDI2050, GDI2648) and *groEL* (GDI2049, GDI2647), which are usually present as single copy in bacteria. In *A. aceti*, over-expression of the *groESL* operon led to augmented resistance to acetic acid [30], which may be explained by the fact that chaperonins protect proteins under denaturing conditions such as low pH [31].

Polysaccharides: CPS, EPS and LPS

Cell-surface components that are commonly involved in plant-bacteria interactions include capsular polysaccharides (CPS), exopolysaccharides (EPS), and lipopolysaccharide (LPS). On the GDI chromosome we found nine CDS related to polysaccharide encapsulation (GDI2398 to GDI2402 and GDI2409 to GDI2413). The GDI genome contains several CDS related to lipopolysaccharide biosynthesis. Five CDS (GDI3265, GDI1647, GDI1652, GDI1447 and GDI0495) encode glycosyltransferases, three CDS (GDI2535, GDI2549 and GDI2493) may be involved in lipopolysaccharide transport, one CDS (GDI2975) encodes an O-antigen polymerase, and there is an ADP-heptose synthase (GDI1133) and a nucleotidyl transferase (GDI0713). Seven CDS (GDI2490, GDI2971, GDI2492, GDI2544, GDI2549, GDI1898 and GDI1899) related to the synthesis of other EPS such as beta-glucans and exooligosaccharides were also identified. These CDS are dispersed over the GDI genome and encode *exoF*, *exoZ*, *exoY*, *exoO*, *exoP*, *exoN* and *exoC*, respectively. Homologs of these CDS are involved in the interaction between rhizobia and their host plants [32]. GDI has a cluster (GDI2535-GDI2552) containing 14 CDSs that is similar to the *gum* cluster of *Azoarcus sp*.BH72, *X. campestris* and *X. fastidiosa*. The *gum* cluster in *X. campestris* is

responsible for the synthesis of an EPS that is involved in host plant colonization and virulence [33]. However, not all genes from the *gum* operon are present in GDI. We found eight CDSs (GDI2552, GDI2549, GDI2547, GDI2538, GDI2550, GDI2535, GDI2542 and GDI2548) which represent the genes *gumB*, *C*, *D*, *E*, *H*, *J*, *K* and *M*, respectively. The genes *gumF*, *G*, *I* and *L* are not present in the GDI genome. As GDI is not virulent, this cluster may be related with colonization and survival. In addition, it is proposed that the viscous nature of EPS helps localize and stabilize hydrolytic enzymes produced by the bacteria [34]. We found a putative endoglucanase protein (GDI2537) in the *gum* cluster that may degrade plant cell walls in order to facilitate the active penetration of the bacteria and thereafter the colonization. In addition, the production of hydrolytic enzymes by GDI has been observed [35].

Biological Nitrogen Fixation (BNF)

The genetics and biochemistry of BNF and nitrogen utilization by *G. diazotrophicus* have been previously investigated to some extent. Corroborating previous studies [36], we have found that the GDI structural genes for nitrogenase *nifHDK* are arranged in a cluster (GDI0425-GDI0454), which also contains other N₂ fixation-related genes, such as *fixABCX*, *modABC* and *nifAB*. Other related genes, *ntrX*, *ntrY* and *ntrC* (GDI2263, GDI2264, and GDI2265) are localized elsewhere in the chromosome in a 5.2 kb cluster. There are three copies of *nifU* homologous genes, one localized in the *nif* cluster (GD0447), and the other two scattered on the GDI chromosome (GDI1392 and GDI3055). No *draT* or *draG* homologs were found in GDI, confirming that nitrogenase activity is not regulated at the post-translational level. It has been suggested that post-translational modulation in *G. diazotrophicus* might be mediated by a FeSII Shethna protein [37], but no such CDS was identified. However, many other FeSII protein genes are present, and they possible candidates for this role. The apparent absence of *nifL* as a *nifA* activity modulator in response to the cell O₂ status in GDI [38] is in agreement with the lack of a *nifL* homolog on the genome. The *nifA* protein appears to be inherently sensitive to O₂. In *G. diazotrophicus*, the main route for assimilation of ammonia is believed to occur through the glutamine synthetase/glutamate synthase pathway (GS/GOGAT encoded by *glnA* and *gltDB*, respectively) [39]. However, the genome analysis suggests the existence of alternative routes, where the putative enzymes NAD-synthase (GDI0919), aminomethyltransferase (GDI2317), histidine ammonia-lyase (GDI0550) and D-amino acid dehydrogenase (GDI2422) would incorporate ammonia into different compounds. The enzymatic activity of GS is known to be regulated by an adenylyltransferase enzyme, which is probably encoded by *glnE* (GDI3425). The glutamate dehydrogenase gene was not found in GDI,

although its activity was demonstrated for *G. diazotrophicus* strain Pal3 [38].

Signaling and quorum sensing

The GDI genome contains 16 GGDEF family genes that are involved in the synthesis of the second messenger cyclic di-GMP, which has been shown to regulate cellulose synthesis and other processes such as transitions between sessile and planktonic lifestyle and pathogenesis [39]. There are three cytoplasmic and 14 membrane-bound histidine kinase signaling proteins, the majority of which form two-component signaling systems with a neighboring response regulator gene. Among these histidine kinases are homologs of the *kdpD* (GDI1566), *envZ* (GDI3079), *chvG* (GDI1265), *ntrY* (GDI2264), *ntrB* (GDI2266) and *phoB* (GDI3817) genes. In addition, there are two adjacent hybrid histidine kinase/response regulator genes that are organized in an apparent operon (GDI3283-3293) that contains several chemotaxis genes and a proteolytic system encoded by *hslUV* that is absent in GOX. Chemotaxis enables microorganisms to move towards beneficial or away from harmful substances in their environments by means of flagellar motility. The *G. diazotrophicus* genome contains nine methyl-accepting proteins (MCPs, chemotaxis sensor proteins), the majority of which have close homologs in rhizobia, but not in the phylogenetically related non-endophyte GOX, which has only three MCP genes [40]. Quorum sensing has been shown to be important in traits such as virulence, biofilm formation and swarming motility in many bacteria [41]. In the *Azoarcus* sp BH72 genome, quorum sensing genes were not found, and it was suggested that this was compatible with a non-pathogenic interaction of *Azoarcus* sp BH72 with the host plant [12]. Nevertheless, GDI, which inhabits a niche similar to *Azoarcus* sp BH72, has three quorum sensing genes: one *luxI* autoinducer synthase gene (GDI2836) and two *luxR*-type transcriptional regulator genes (GDI2837, GDI2838). Quorum sensing genes are also present in several rhizobial genomes, and they play roles in nodulation and nitrogen fixation [42].

Plant Growth-Promoting (PGP) Traits

There are several indications that GDI promotes plant growth by more than a few independent mechanisms besides nitrogen fixation, including synthesis of phytohormones and increased uptake of nutrients [43]. Recent work has shown that mutations in two genes involved in cytochrome c biogenesis reduced auxin levels to 10% of the wild-type strain [44], suggesting their involvement in indole acetic acid (IAA) production, and indicating that GDI has at least two independent pathways for auxin biosynthesis. In addition, characterization of the IAA biosynthetic route in GDI has shown that auxin is mostly synthesized by the Indole-3-pyruvic acid (IPyA) pathway [44]. Although no CDS encoding an indole 3-pyruvate

carboxylase was found in GDI genome, we cannot rule out that the biochemical activity could be executed by one of the many putative decarboxylases identified in the genome. The presence of genes encoding enzymes such as aromatic-L-amino-acid decarboxylase (GDI1891), amine oxidase (GDI1716) and aldehyde dehydrogenases (GDI0311, GDI0461, GDI640) suggests that the bacteria might synthesize IAA via the tryptamide pathway (TAM). Also, the presence of two genes coding for putative nitrilases (GDI0018, GDI3743) suggests that IAA might be produced by the indole-3-acetonitrile pathway (IAN). In addition to phytohormone production, some rhizosphere-associated bacteria can stimulate plant growth by secreting a mixture of plant volatiles, mainly 3-hydroxy-2-butanone (acetoin) and 2,3-butanediol [45]. Although the role of GDI in PGP has been studied, no attention has been paid to the production of volatiles. We found GDI is likely to be capable to synthesize acetoin once the genome sequence encodes two enzymes of the pathway; acetolactate synthase (GDI0022, GDI0023) and acetoin diacetyl reductase (GDI2623). In addition, although an acetolactate decarboxylase has not been identified, 2-acetolactate can be converted to diacetyl spontaneously in the presence of oxygen [46]. It has been shown for *Azospirillum brasilense* that the production and secretion of polyamines promote plant growth [47]. The presence of genes coding for enzymes for the synthesis (GDI0476, GDI2322) and secretion (GDI2595) of spermidine in the *G. diazotrophicus* Pal5 genome sequence suggests that this polyamine may also contribute to PGP. *G. diazotrophicus* has been shown to synthesize the gibberellins GA1 and GA3 [48]. Although the gibberellin biosynthesis machinery in bacteria is largely unknown, recent studies have suggested likely biosynthetic mechanisms in *Bradyrhizobium japonicum* [49]. The GDI genome contains genes for the synthesis of the diterpenoid precursor isopentenyl diphosphate through the non-mevalonate pathway. Condensation reactions of this precursor to form geranylgeranyl diphosphate may be performed by the geranyltranstransferase *ispA* (GDI1861). However, homologs of the genes responsible for the cyclization of geranylgeranyl diphosphate in *B. japonicum* (ent-copalyl diphosphate and ent-kaurene synthase) are apparently absent in the GDI genome and therefore the mechanism of cyclization of geranylgeranyl diphosphate to ent-kaurene remains unknown. However, a putative squalene cyclase (GDI1620) could fulfill such function, since a study with recombinant squalene cyclase has shown some cyclization of geranylgeraniol by this enzyme [50]. Oxidation steps of ent-kaurene, necessary to produce GA₁ and GA₃, may be catalyzed by two cytochromes P450 (GDI2364 and GDI2593), homologs of which are absent in other acetobacteraceae genomes, thus suggesting a likely specific role in *G. diazotrophicus*. It has been reported that the capacity of *G. diazotrophicus* to antagonize diverse plant pathogens such as fungi, and

bacteria contributes to increasing its ability to survive under environmental stress and leads to an improvement in plant fitness which may have important consequences for agricultural productivity [9,10]. Its genome sequence encodes a large repertoire of genes whose products oppose attack from competing microbes, such as drug efflux systems, and acriflavin and fusaric acid resistance proteins. On the other hand, GDI may also produce a broad variety of proteins such as lytic enzymes and phospholipases and antibiotic biosynthetic pathways that could be toxic to other organisms. The secretion of a lysozyme-like bacteriocin by *G. diazotrophicus*, for instance, inhibits *Xanthomonas albilineans* growth [9]. Indeed, GDI encodes a putative lysozyme-like bacteriocin (GDI0416 and GDI0415).

Sugar metabolism and energy generation

Sucrose is the common carbon source used for isolation of *Gluconacetobacter diazotrophicus* from sugarcane and other plants in the semi-solid LGIP medium [51]. However, sucrose is not directly metabolized by the bacteria. Experimental evidence has shown that there is a constitutively expressed levansucrase (LsdA GDI0471), secreted to the periplasm via a specific signal peptide-dependent pathway, that converts sucrose to beta-1,2 -oligofructans and levan [52]. In addition, a fructose-releasing exo-levanase (LsdB GDI 0477) probably controlled by an antitermination inducer system converts polyfructans into fructose [53]. A type II secretion operon (GDI481-GDI 490) is required for the transport of LsdA across the outer membrane [54]. The transport of LsdB to the periplasm involves the cleavage of the N-terminal peptide signal, and it is induced during growth of the bacteria with low fructose levels but repressed by glucose [55].

In *G. diazotrophicus* oxidation of glucose to gluconate in the periplasmic space is the first step in glucose metabolism by GDI [56]. Gluconate may be synthesized by the product of three CDS encoding membrane-bound quinoprotein glucose dehydrogenases (GDI3277, GDI0325 and GDI0539) in accordance with the observed high activity of PQQ-GDH detected in glucose-containing batch culture of GDI strain Pal3 grown mainly under biological nitrogen fixation and/or C-limitation conditions [57]. A NAD-GDH (GDI2625) also participates in the glucose oxidation (intracellularly) when glucose is in excess [57]. Further periplasmic oxidation of gluconate to 2-ketogluconic acid occurs by a putative three-subunit flavin-dependent gluconate-2-dehydrogenase (GDI0854, GDI0855 and GDI0856). Gluconate dehydrogenases (extracellular, dye-linked and intracellular, NAD-Linked) activities have been demonstrated in GDI strain Pal3 grown in presence of gluconate with 2-ketogluconate the major compound accumulated (57). The production of 5-ketogluconate and 2,5 di-ketogluconate are probably

mediated by a glucose/methanol/choline oxidoreductase (GDI0859) and a putative alcohol dehydrogenase cytochrome c/gluconate 2-dehydrogenase acceptor (GDI0860). High activities of 2-ketogluconate reductase (NAD linked) have been detected in a GDI Pal3 strain grown with gluconate [58].

CDS for transport (GDI3258) and phosphorylation (GDI0293) proteins indicate that gluconate can also be directly driven into the pentose phosphate route (PPP), supporting the experimental data [58]. The presence of a kinase (GDI3115), a 2-ketogluconate reductase (GDI 3432) and a 6-phosphogluconate dehydrogenase-NAD (GDI2166) corroborates with the experimental data which shows that the PPP is the main C-metabolism route in GDI following the oxidation of glucose to gluconate [57].

Different from GOX, CDSs encoding a complete respiratory chain complex I (*nuoA - nuoN* or complex I proton-translocating NADH-quinone oxidoreductase; GDI2459-GDI2471) are present in the GDI genome [59]. The GDI genome contains CDS that encode L-sorbose dehydrogenases (GDI0574 and GDI3764), membrane-bound small and large subunits (GDI3280 and GDI3281) and the cytochrome c subunit (GDI3279) of aldehyde dehydrogenase, indicating that GDI may be able to synthesize the industrially important substances such as L-ascorbic acid (vitamin C) and its precursor 2-keto-L-gulonic acid [60].

Genome Features in Accessory Regions

Type IV secretion system

Type IV secretion systems (T4SS) are multi-subunit cell envelope-spanning structures, ancestrally related to bacterial conjugation machines, that transfer proteins, DNA and nucleoprotein complexes across membranes [61]. Moreover, T4SSs have been described as essential pathogenicity factors and recently it has been indicated that TSS4 can also increase host adaptability in *Bartonella sp.* [22]. GDI has 4 complete T4SS in the chromosome which are similar to bacterial conjugation machines (*trb*) of *Agrobacterium tumefaciens* [62] and Ti (tumor inducing) *Enterobacter* IncP plasmid R751 [63]. Although the order of the *trb* genes in the operon is conserved (*trbB*, -C, -D, -E, -J, -L, -F, -G, -I), two genes are missing from the original *trb* operon (*trbK* and *trbH*). The gene *trbK* has been reported as non-essential but *trbH* has been reported as essential for conjugal transfer of *Agrobacterium* tumor inducing plasmid pTiC58 [63]. Another difference is that, in *Agrobacterium tumefaciens* and *Enterobacter* IncP plasmid R751, the first gene in the operon is *traI*, which is an essential signal for the quorum-sensing regulation of the Ti plasmid conjugation transfer [64]. In GDI the first gene in the operon is *traG*, which is essential for DNA transfer

in bacterial conjugation. This gene is thought to mediate interactions between the DNA-processing (Dtr) and the mating pair-formation (Mpf) systems [65]. T4SS have been found in many different organisms [66], from pathogenic to mutualistic endosymbiont organisms (for instance, *Helicobacter pylori*, *Legionella pneumophila*, *Brucella spp.*, *Bartonella spp.*, *Rickettsia spp.*, *Coxiella spp.*, *Anaplasma marginale*, *Ehrlichia spp.*, *Agrobacterium tumefaciens*, *Wolbachia spp.*). All four complete T4SS operons in the GDI chromosome were found in accessory regions (GI4, GI12, and twice in GI21), suggesting that the bacteria acquired the ability to translocate macromolecules across the cell envelope to the plant. The four copies of the T4SS operon diverge by the presence of a variable region between the *traG* and the *trbB* genes that include transcriptional regulators *mucR* and *araC*, a DNA-binding protein HU-beta, an aldo/keto reductase and hypothetical proteins. These genes might confer specific functions to each T4SS copy.

Flagella and pili

In many organisms, flagella are involved in motility, adherence, biofilm formation and host colonization [67]. GDI has a large accessory region (GI15) with at least 40 genes predicted to encode functions related to motility. This observation is in accordance with the presence of peritrichous flagella on the GDI cell surface. Next to the motility cluster there is a putative *tad* locus (*Flp-1*, *cpaABC*, *cpaEF*, and *tadBCDG*) which probably encodes the machinery for the synthesis of Flp (fimbrial low-molecular-weight protein) pili, which form a subfamily in the type IVb pilus family. In *Actinobacillus*, *Haemophilus*, *Pasteurella*, *Pseudomonas*, *Yersinia* and *Caulobacter*. Flp pili are essential for biofilm formation, colonization and pathogenesis [68]. Additionally, several pseudopilins (GDI0483, GDI0484, and GDI0485) were identified as part of a type II secretion system. Recently, it has been shown that flagella-less mutant of GDI was non-motile and displayed reduced capacity to form biofilm [69]. These findings suggest that these genes were acquired by HGT and play an important role in the interaction with the plant.

Conclusion

Despite the potential impact of endophytes on the environment and on crop production, our current knowledge of their biology is limited. Analysis of the *G. diazotrophicus* PAL5 complete genome sequence provides important insights into the endophytic relationship, and suggests many interesting candidate genes for post-genomic experiments.

The genome reveals an unexpectedly high number of mobile elements for an endophytic bacterium; it is in fact the endophyte with the highest frequency of mobile genes per Mb of genome. The high number of mobile elements

seems to be associated with a high number of HGT events. The analysis of HGT shows that most of the genes are more similar to genes from the order rhizobiales (40%), suggesting that a likely previous niche was located in the rhizosphere. Thus, a recent evolutionary bottleneck and consequent relaxation of selection, due to a possible change of niche, is probably the hypothesis that could best explain the high number of HGT [15].

In addition, to change niche from rhizosphere to endophytic, the bacteria should have features that would allow it to penetrate the plant. The putative *gum*-like cluster containing an endoglucanase could be important in this regard. Moreover, the limited similarity with the *gum*-like cluster from *X. campestris* and the absence of some genes found in *X. campestris* may mean that the cluster adapted to a non-virulence profile. However, the ability to penetrate the plant is not enough to transform it into an endophyte; the bacteria must evolve together with the plant to create a more depended relationship. The genome has many features to enhance plant fitness such as BNF, phytohormones and biocontrol genes, and all of them lie in the core of the genome or have a very low "Alien score". We propose that these features were important to create a dependent relationship, and may have helped GDI to spread out and occupy this niche. In contrast, many features that may be related to bacteria-plant interaction are found in genome islands, including type IV secretion systems, flagella, pili, chemotaxis, biofilm, capsular polysaccharide and some transport proteins. The overall result suggests that it is more likely that GDI acquired many features that are important for an endophytic lifestyle. Thus, experimental analyses of genes from genome islands may reveal an important source of gene candidates that will enhance our understanding of bacteria-plant relationship.

Finally, comparison of genome sequences of *Gluconacetobacter diazotrophicus* and *Azoarcus sp.* BH72 shows that these endophytic diazotrophic bacteria adopted very different strategies to colonize plants. A limited number of genomic features, such as the large number of TonB receptors, the *gum*-like and *nif* clusters, and osmotolerance mechanisms are common to both endophytic diazotrophic bacteria. On the other hand, *Gluconacetobacter diazotrophicus* has a larger number of transport systems, and it is capable of growing on a wide variety of carbon sources, while *Azoarcus sp.* BH72 has rather complex signaling mechanisms to communicate with its plant host.

Methods

Strain

Gluconacetobacter diazotrophicus strain PAL 5 (type strain) was isolated from sugarcane roots collected in Alagoas Sate, Brazil using the nitrogen-free semi-solid LGIP medium [2]. It was deposited at the Embrapa Agrobio-

logia Culture Collection and received the identification number BR 11281 (BR-stands for the Brazilian Nitrogen-fixing bacteria Culture Collection). Later on, this strain was deposited by Johanna Dobereiner at the American Type Culture Collection (ATCC 49037) and also at the Culture Collection Laboratorium von Microbiologie, Belgium (LMG 7603) [70].

Genome sequencing, assembly and annotation

All the libraries were prepared with total bulk DNA originated from a Pal 5 lyophilized tube culture provided by the Embrapa Agrobiologia Culture Collection. Pal5 was grown in 500 mL Erlenmeyer flasks containing 200 mL of DYGS medium (Rodrigues-Neto et al., 1986) during 48 h at 200 rpm and 30°C. DNA extraction was performed according with the CTAB method [71]. Phenol: chloroform: iso-amilic alcohol (25:24:1) and chloroform: iso-amilic alcohol (24:1) washing steps were repeated 2 times to guarantee removal of cells debris and other contaminants during DNA extraction.

DNA shotgun libraries with insert sizes of 0.5-1 kb, 2-3 kb and 4-6 kb were constructed in pUC18 vectors and 10-17 kb in the cosmid pLARF3. Plasmid clones were end-sequenced on ABI377 and ABI3100 (Applied Biosystems) and MegaBACE 1000 (GE Healthcare) sequencers. A total of 103,506 high-quality reads were obtained and assembled into contigs using the Phrap assembly tool. For gap closure, 16,963 additional reads were obtained through PCR direct sequencing and primer-walking on plasmids. Manual editing was done using the GAP4 software package [72]. Genome integrity was verified by a physical map constructed using PFGE and hybridization with 42 single-copy and rDNA probes [73]. Initial automatic gene prediction was done using GLIMMER [74], and subsequently manually curated with reference to codon-specific positional base preferences. Before the manual annotation of each predicted gene, different tools were used. Similarity search was performed against different databases including Uniprot [75], PROSITE [76], nr, Pfam [77], and InterPro [78]. Additionally, SignalP [79], TMHMM [80] and tRNAscan-SE [81] were applied. All the data were viewed within the Artemis [82] program where the function of each gene was manually curated.

Annotation colors

Pathogenicity/Adaptation/Chaperones, dark blue; Energy metabolism (glycolysis, electron transport etc.), gray; Information transfer (transcription/translation, DNA/RNA modification), red; Surface structures (IM, OM, secreted, LPS)), green; Stable RNA, cyan; Degradation of large molecules, light blue; Degradation of small molecules, purple; Central/intermediary/miscellaneous metabolism, yellow; Unknown and conserved hypothetical, orange; Regulators, magenta; Pseudogenes and partial

genes, black; Phage/IS elements, pink; miscellaneous information (e.g. Prosite but no function), brown.

Nucleotide sequence accession numbers

The genomic sequence reported in this article has been deposited in the EMBL database under accession numbers [AM889285](http://www.ebi.ac.uk/EMBL/seq_data/embl/seq_data/AM889285), [AM889286](http://www.ebi.ac.uk/EMBL/seq_data/embl/seq_data/AM889286) and [AM889287](http://www.ebi.ac.uk/EMBL/seq_data/embl/seq_data/AM889287). The genome annotation and features are available at <http://www.bioqmed.ufri.br/bertalan/>.

Core and accessory regions

The core regions were determined by quartets analysis (quartets of orthologous proteins), using reciprocal best hit of Blastp. The accessory regions were determined by a combination of two different methods: GC3 and IVOMs. The GC3 analyzes the percent of GC in the third base of the codon in each gene. For both methods, the regions indicated as accessory genes were manually checked for integrases, tRNAs and repeats (direct and inverted). The beginning and end of each the accessory region were defined by both methods and, in the case of bacteriophages, the genome islands were extended when evidence of the insertion point was found.

Reciprocal Best Hits

Reciprocal best hits comparison was done using only the complete bacterial genomes publicly available at <ftp://ftp.ncbi.nih.gov/genomes/Bacteria/all.faa.tar.gz>. Only reciprocal best hits with identity greater of 30% and alignment greater than 70% were selected.

Plant Endophyte comparison

Six complete endophyte genomes were used to represent the endophyte group and three closest complete genomes phylogenetically to GDI were used to represent the core genome. Endophyte genomes were *Azoarcus* sp. BH72, *Burkholderia phytofirmans* PsJN, *Enterobacter* sp. 638, *Methylobacterium populi* BJ001, *Pseudomonas putida* W619 and *Serratia proteamaculans* 568. Core genome species were *Acidiphilium cryptum* JF-5, *Gluconobacter oxydans* 621H and *Granulibacter bethesdensis* CGDNIH. Only reciprocal best hits with more than 30% identity and 70% alignment were accepted.

Abbreviations

BNF: Biological Nitrogen Fixation; GDI: *Gluconacetobacter diazotrophicus* PAL5; HGT: Horizontal Gene Transfer; GI: Genome Island; CDS: Coding Sequences; PHX: Predicted Highly Expressed Genes; T4SS: Type IV secretion system; ACC: *Acidiphilium cryptum* JF-5; GOX: *Gluconobacter oxydans* 621H; GRB: *Granulibacter bethesdensis* CGDNIH; GC3: G+C content of synonymous third position; IVOMs: Interpolated Variable Order Motifs; IS: Insertion sequence; BBH: Blast Best Hit; Flp: fimbrial low-molecu-

lar-weight protein; bp: base pairs; Dtr: DNA-processing; Mpf: mating pair formation; RBH: Reciprocal Best Hits.

Authors' contributions

PF coordinated the study. MB, AP, DC, TU, WK, PB, LP, DF, HM, JJ, VF, BF, AB, ES, GA, AN, RA, DC, DO, TS, JM, AV, PG, EN, VP, JA, AO, LR, HG, ER, JB, LF, VM, RS, ML and PF conducted genome sequencing. MB and ML conducted genome assembly. MB, LL, PB, SR, GF, VF, MF, GM, AC, AN, RA, JM, HG, EN, VP, ML, LS, JM, KT, JA, MV, SS, AO, LR, JB, CM, AH, SA, AMC, WA, MD, RS, EO, AR, ML, OM and PF performed sequence annotation. MB performed bioinformatics analyses and comparative genome analyzes. MB, SS, VP, EN, AO, LR, JB, CR, AMC, and PF analyzed the results and participated in writing sections of the manuscript. MB and PF assembled and wrote the final version of the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

Distribution of mobile elements in plant endophyte complete genomes.

The percentage column: Percentage of total number of mobile elements from all CDS annotated on the endophyte complete genomes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S1.PDF>]

Additional file 2

Predicted Highly Expressed (PHX) genes. The PHX and proteomic analysis was used to indicate potentially important genes in the GDI genome.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S2.XLS>]

Additional file 3

16S phylogenetic tree from Alphaproteobacteria. The Neighbor joining phylogenetic tree of 16S from Alphaproteobacteria was done using ClustalX. In blue are the three completed genomes closest to G. diazotrophicus PAL5 available in GenBank.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S3.EPS>]

Additional file 4

The 28 genome islands (GI) identified by GC3 and IVOMs. The GI column has the ID for each genome island. The integrase column shows which kind of integrase was found in each genome island. The CDS column shows how many CDS are inside the genome island. The Alien+GC3 column show how many CDS in each genome island were identified as accessory by both methods. The Related column shows which kinds of genes were found in each genome island.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S4.PDF>]

Additional file 5

Variation in G. diazotrophicus strains. 20 different strains were tested for gene variation. 37 CDS were selected from 21 putative genome islands and 17 CDS were selected from putative core regions of the chromosome as control. (+): PCR positive. (-): PCR negative.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S5.PDF>]

Additional file 6

Presence of homologues of the trbE gene among G. diazotrophicus strains. Total DNA of 11 Gluconacetobacter strains was completely digested with restriction enzymes EcoRI (a.) or EcoRV (b.), separated on agarose gel and submitted to Southern blot analysis using a fragment of CDS GDI0133 (trbE, part of type IV secretion system) as a probe. Numbers 1-10 represent G. diazotrophicus strains: Pal5 (1), 3R2 (2), URU (3), 38f2 (4), PRI50(5), Pal3 (6), AF3 (7), PCRI (8), PPe4 (9), CNFe-550 (10). Number 11 represents G. johannae. In strain Pal5, only 3 bands are present, although the genome sequence indicates the presence of four copies of the trbE gene. However, the fourth trbE paralog (GDI1016) is more dissimilar to the probe sequence than the other three (GDI0133, GDI2742 e GDI2911), which may have prevented hybridization.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S6.JPEG>]

Additional file 7

Distribution of percent ID from RBH results. In red all the RBH from Rhodospirillales order. In yellow all RBH from other Alphaproteobacteria class and in blue RBH from other genomes beside Alphaproteobacteria class.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S7.EPS>]

Additional file 8

Number of Reciprocal Best Hits (RBH) in accessory and core regions.

The first column shows the number of RBH for each organism in parentheses. The RBH in columns show the total number of RBH for each organism. The RBH % by organism columns shows the percent of RBH in relation with the total number of RBH found in accessory and core regions. RBH result has 708 RBH in accessory regions and 2,258 in core regions. The RBH % by organism columns shows the percentage of RBH in accessory and core regions for each organism or group. GOX = Gluconobacter oxydans 621H, GBE = Granulibacter bethesdensis CGD-NIH, ACR = Acidiphilium cryptum JF-5, Rhiz = All the complete genomes from Rhizobiales order, Other Alpha = All other complete genomes from Alphaproteobacteria class, Beta = All complete genomes from Betaproteobacteria class, Gamma = All complete genomes from Gammaproteobacteria class, Others = All other complete genomes.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S8.PDF>]

Additional file 9

Endophyte comparison gene list. Endophyte gene list.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S9.XLS>]

Additional file 10

Endophyte comparison. In gray, genes similar to all genomes (core + endophyte, see Methods). In blue, genes present in all endophyte but not in core genomes. In red, genes only similar to GDI and *Azoarcus* sp BH72. In purple, genes only similar to GDI and *Methylobacterium populi* BJ001 and in green genes that are only present in GDI and at least two other endophyte genome.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S10.EPS>]

Additional file 11

Comparison of main signalling protein categories. AT, *Agrobacterium tumefaciens* C58; BJ, *Bradyrhizobium japonicum* USDA110; ML, *Mesorhizobium loti* MAFF303099, SM, *Sinorhizobium meliloti* 1021, GO, *Gluconobacter oxydans* 621H; RP, *Rickettsia prowazekii* MadridE; AB, *Azoarcus* sp. BH72; AE, *Azoarcus* sp. EbN1; XF, *Xylella fastidiosa* 9a5c; EC, *Escherichia coli* K12-MG1655.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S11.PDF>]

Additional file 12

Comparison of main transport-related protein categories. AT, *Agrobacterium tumefaciens* C58; BJ, *Bradyrhizobium japonicum* USDA110; ML, *Mesorhizobium loti* MAFF303099, SM, *Sinorhizobium meliloti* 1021, GO, *Gluconobacter oxydans* 621H; RP, *Rickettsia prowazekii* MadridE; AB, *Azoarcus* sp. BH72; AE, *Azoarcus* sp. BH72, XF, *Xylella fastidiosa* 9a5c; EC, *Escherichia coli* K12-MG1655.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S12.PDF>]

Additional file 13

Comparison among the two *Gluconacetobacter diazotrophicus* Pal5 genomic sequences. GDI-BR, NCBI RefSeq NC_010125, GDI-US, NCBI RefSeq NC_011365, GIs, Genome Islands.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S13.PDF>]

Additional file 14

CDS list of the two *Gluconacetobacter diazotrophicus* Pal5 sequences. Sheet 1: Blast best hits list of CDS found in both genomes. Sheet 2: List of unique CDS found in chromosome from GeneBank file CP001189. Sheet 3: List of unique genes found in chromosome from GeneBank file AM889285 (this work).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-10-450-S14.XLS>]

Acknowledgements

This work is dedicated to the memory of Johanna Döbereiner. This work was funded with grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro Carlos Chagas Filho (FAPERJ) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). We are grateful to Julian Parkhill and Martha Sorenson for helpful discussions and for critical reading of the manuscript.

References

- Boddey RM, Döbereiner J: **Nitrogen fixation associated with grasses and cereals: recent progress and perspectives for the future.** *Fertilizer Res* 1995, **42**:241-250.
- Cavalcante VA, Döbereiner J: **A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane.** *Plant Soil* 1988, **100**:23-31.
- James EK, Reis VM, Olivares FL, Baldani JJ, Döbereiner J: **Infection of sugarcane by the nitrogen-fixing bacterium *Acetobacter diazotrophicus*.** *J Exp Bot* 1994, **45**:757-766.
- Muthukumarasamy R, Cleenwerck I, Revathi G, Vadivelu M, Janssens D, Hoste B, Gum KU, Park KD, Son CY, Sa T, Caballero-Mellado : **Natural association of *Gluconacetobacter diazotrophicus* and diazotrophic *Acetobacter peroxydans* with wetland rice.** *J Syst Appl Microbiol* 2005, **283**:277-286.
- Boddey RM, Urquiaga S, Reis VM, Döbereiner J: **Biological nitrogen fixation associated with sugarcane.** *Plant Soil* 1991, **137**:111-117.
- Baldani JJ, Caruso LV, Baldani VLD, Goi SR, Döbereiner J: **Recent advances in BNF with non-legume plants.** *Soil Biol Biochem* 1997, **29**:911-922.
- Dong Z, Canny MJ, McCully ME, Robredo MR, Cabadilla CF, Ortega E, Rodes R: **A Nitrogen-Fixing Endophyte of Sugarcane Stems (A New Role for the Apoplast).** *Plant Physiology* 1994, **105**:1139-1147.
- Sevilla M, Burris RH, Gunapala N, Kennedy C: **Comparison of benefit to sugarcane plant growth and $^{15}\text{N}_2$ incorporation following inoculation of sterile plants with *Acetobacter diazotrophicus* wildtype and Nif- mutant strains.** *Mol Plant Microbe Interact* 2001, **3**:358-366.
- Blanco Y, Blanch M, Pin D, Legaz ME, Vicente C: **Antagonism of *Gluconacetobacter diazotrophicus* (a sugarcane endosymbiont) against *Xanthomonas albilineans* (pathogen) studied in alginate-immobilized sugarcane stalk tissues.** *J Biosci Bio eng* 2005, **4**:366-371.
- Mehnaz S, Lazarovits G: **Inoculation effects of *Pseudomonas putida*, *Gluconacetobacter azotocaptans*, and *Azospirillum lipoferum* on corn plant growth under greenhouse conditions.** *Microb Ecol* 2006, **3**:326-335.
- Saravanan VS, Kalaiarasan P, Madhaiyan M, Thangaraju M: **Solubilization of insoluble zinc compounds by *Gluconacetobacter diazotrophicus* and the detrimental action of zinc ion (Zn^{2+}) and zinc chelates on root knot nematode *Meloidogyne incognita*.** *Lett Appl Microbiol* 2007, **3**:235-241.
- Krause A, Ramakumar A, Bartels D, Battistoni F, Bekel T, Boch J, Boehm M, Friedrich F, Hurek T, Krause L, et al.: **Complete genome of the mutualistic, N_2 -fixing grass endophyte *Azoarcus* sp. strain BH72.** *Nat Biotechnol* 2008, **24**:1385-1391.
- Fouts DE, Tyler HL, De Boy RT, Daugherty S, Ren Q, Badger JH, Durkin AS, Huot H, Shrivastava S, Kothari S, et al.: **Complete genome sequence of the N_2 -fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice.** *PLoS Genet* 2008, **4**:e1000141.
- Hapmap: **Endophyte complete bacterial genomes** [<http://www.expasy.ch/sprot/hamap/interactions.html>]
- Parkhill J, Sebahia M, Preston A, Murphy LD, Thomson N, Harris DE, Holden MT, Churcher CM, Bentley SD, Mungall KL, et al.: **Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*.** *Nat Genet* 2003, **35**:32-40.
- Karlin S, Barnett MJ, Campbell AM, Fisher RF, Mrzcek J: **Predicting gene expression levels from codon biases in alpha-proteobacterial genomes.** *Proc Natl Acad Sci* 2003, **100**:7313-7318.
- Lery LM, Coelho A, von Kruger WM, Gonçalves MS, Santos MF, Valente RH, Santos EO, Rocha SL, Perales J, Domont GB, et al.: **Protein expression profile of *Gluconacetobacter diazotrophicus* PAL5, a sugarcane endophytic plant growth promoting bacterium.** *Proteomics* 2008, **8**:1631-1644.
- Wernegreen JJ: **Genome evolution in bacterial endosymbionts of insects.** *Nat Rev Genet* 2002, **3**:850-861.
- Zhaxybayeva O, Gogarten JP: **Bootstrap, and Bayesian probability and maximum likelihood mapping: exploring new tools for comparative genome analyses.** *BMC Genomics* 2002, **3**:4-19.
- Vernikos GS, Parkhill J: **Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands.** *Bioinformatics* 2006, **22**:2196-2203.

21. Ochman H, Lawrence JG, Groisman EA: **Lateral gene transfer and the nature of bacterial innovation.** *Nature* 2000, **405**:299-304.
22. Saenz HL, Engel P, Stoeckli MC, Lanz C, Raddatz G, Vayssier-Taussat M, Birtles R, Schuster SC, Dehio C: **Genomic analysis of *Bartonella* identifies type IV secretion systems as host adaptability factors.** *Nat Genet* 2007, **39**:1469-1476.
23. Moreno-Hagelsieb G, Latimer K: **Choosing BLAST options for better detection of orthologs as reciprocal best hits.** *Bioinformatics* 2008, **24**:319-324.
24. Reis VM, Döbereiner J: **Effect of high sugar concentration on nitrogenase activity of *Acetobacter diazotrophicus*.** *Arch Microbiol* 1998, **171**:13-18.
25. Epstein W: **The roles and regulation of potassium in bacteria.** *Prog Nucleic Acid Res Mol Biol* 2003, **75**:293-320.
26. Wood JM: **Bacterial osmosensing transporters.** *Methods Enzymol* 2007, **428**:77-107.
27. Tejera NA, Ortega E, González-López J, Lluich C: **Effect of some abiotic factors on the biological activity of *Gluconacetobacter diazotrophicus*.** *J Appl Microbiol* 2003, **95**:528-535.
28. Mullins EA, Francois JA, Kappock TJ: **A specialized citric acid cycle requiring succinyl-coenzyme A (CoA):acetate CoA-transferase (AarC) confers acetic acid resistance on the acidophile *Acetobacter acetii*.** *J Bacteriol* 2008, **190**:4933-4940.
29. Nakano S, Fukaya M, Horinouchi S: **Putative ABC transporter responsible for acetic acid resistance in *Acetobacter acetii*.** *Appl Environ Microbiol* 2006, **72**:497-505.
30. Okamoto-Kainuma A, Yan W, Kadono S, Tayama K, Koizumi Y, Yanagida F: **Cloning and characterization of groESL operon in *Acetobacter acetii*.** *J Biosci Bioeng* 2002, **94**:140-147.
31. Nakano S, Fukaya M: **Analysis of proteins responsive to acetic acid in *Acetobacter*: molecular mechanisms conferring acetic acid resistance in acetic acid bacteria.** *Int J Food Microbiol* 2008, **125**:54-59.
32. Skorupska A, Janczarek M, Marczak M, Mazur A, Król J: **Rhizobial exopolysaccharides: genetic control and symbiotic functions.** *Microb Cell Fact* 2006, **16**:5-7.
33. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Puhler A, Ielpi L: ***Xanthomonas campestris* pv *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence.** *J Bacteriol* 1998, **180**:1607-1617.
34. Roper MC, Greve LC, Warren JG, Labavitch JM, Kirkpatrick BC: ***Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines.** *Mol Plant Microbe Interact* 2007, **20**:411-419.
35. Adriano-Anayal M, Salvador-Figueroa M, A OJ, García-Romera I: **Plant cell-wall degrading hydrolytic enzymes of *Gluconacetobacter diazotrophicus*.** *Symbiosis* 2005, **40**:151-156.
36. Lee S, Reth A, Meletzus D, Sevilla M, Kennedy C: **Characterization of a major cluster of *nif*, and *fix*, and associated genes in a sugarcane endophyte, and *Acetobacter diazotrophicus*.** *J Bacteriol* 2000, **182**:7088-7091.
37. Ureta A, Nordlund S: **Evidence for conformational protection of nitrogenase against oxygen in *Gluconacetobacter diazotrophicus* by a putative FeSII protein.** *J Bacteriol* 2002, **184**:5805-5809.
38. Perlova O, Nawroth R, Zellermann EM, Meletzus D: **Isolation and characterization of the *glnD* gene of *Gluconacetobacter diazotrophicus*, encoding a putative uridylyltransferase/uridylyl-removing enzyme.** *Gene* 2002, **297**:159-168.
39. Dow JM, Fouhy Y, Lucey JF, Ryan RP: **The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants.** *Mol Plant Microbe Interact* 2006, **19**:1378-1384.
40. Prust C, Hoffmeister M, Liesegang H, Wiezer A, Fricke WF, Ehrenreich A, Gottschalk G, Deppenmeier U: **Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*.** *Nat Biotechnol* 2005, **23**:195-200.
41. Williams P, Winzer K, Chan WC, Cámara M, Philos Trans R: **Look who's talking: communication and quorum sensing in the bacterial world.** *Soc Lond B Biol Sci* 2007, **362**:1119-1134.
42. Daniels R, De Vos DE, Desair J, Raedschelders G, Luyten E, Rosemeyer V, Verreth C, Schoeters E, Vanderleyden J, Michiels J: **The cin quorum sensing locus of *Rhizobium etli* CNPAF512 affects growth and symbiotic nitrogen fixation.** *J Biol Chem* 2002, **277**:462-468.
43. Saravanan VS, Madhaiyan M, Osborne J, Thangaraju M, Sa TM: **Ecological occurrence of *Gluconacetobacter diazotrophicus* and nitrogen-fixing *Acetobacteraceae* members: their possible role in plant growth promotion.** *Microb Ecol* 2008, **55**:130-140.
44. Lee S, Flores-Encarnación M, Contreras-Zentella M, Garcia-Flores L, Escamilla JE, Kennedy C: **Indole-3-acetic acid biosynthesis is deficient in *Gluconacetobacter diazotrophicus* strains with mutations in cytochrome *c* biogenesis genes.** *J Bacteriol* 2004, **186**:5384-5391.
45. Ryu CM, Farag MA, Hu CH, Reddy MS, Kloepper JW, Paré PW: **Bacterial volatiles induce systemic resistance in *Arabidopsis*.** *Plant Physiol* 2004, **134**:1017-1026.
46. Carballo J, Martin R, Bernardo A, Gonzalez J: **Purification, characterization and some properties of diacetyl(acetoin) reductase from *Enterobacter aerogenes*.** *Eur J Biochem* 1991, **198**:327-332.
47. Perrig D, Boiero ML, Masciarelli OA, Penna C, Ruiz OA, Cassán FD, Luna MV: **Plant-growth-promoting compounds produced by two agronomically important strains of *Azospirillum brasilense*, and implications for inoculant formulation.** *Appl Microbiol Biotechnol* 2007, **75**:1143-1150.
48. Bastian F, Cohen A, Piccoli P, Luna V, Baraldi R, Bottini R: **Production of indole-3-acetic acid and gibberellins A(1) and A(3) by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media.** *Plant Growth Regulation* 1998, **24**:7-11.
49. Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ: **Gibberellin biosynthesis in bacteria: separate ent-copalyl diphosphate and ent-kaurene synthases in *Bradyrhizobium japonicum*.** *FEBS Lett* 2009, **583**:475-480.
50. Hoshino T, Kumai Y, Kudo I, Nakano S, Ohashi S: **Enzymatic cyclization reactions of geraniol, farnesol and geranylgeraniol, and those of truncated squalene analogs having C20 and C25 by recombinant squalene cyclase.** *Org Biomol Chem* 2004, **2**:2650-2657.
51. Baldani JJ, Baldani VLD: **History on the biological nitrogen fixation research in gramminaceous plants: special emphasis on the Brazilian experience.** *Anais da Academia Brasileira de Ciências* 2005, **77**:549-579.
52. Hernandez L, Arrieta J, Menendez C, Vazquez R, Coego A, Suarez V, Selman G, Petit-Glatron MF, Chambert R: **Isolation and enzymatic properties of levansucrase secreted by *Acetobacter diazotrophicus* SRT4, a bacterium associated with sugar cane.** *Biochem J* 1995, **309**:113-118.
53. Menéndez C, Hernández L, Banguela A, País J: **Functional production and secretion of the *Gluconacetobacter diazotrophicus* fructose-releasing exo-levanase (LsdB) in *Pichia pastoris*.** *Enz Microbiol Technol* 2004, **34**:446-452.
54. Arrieta JG, Sotolongo M, Menéndez C, Alfonso D, Trujillo LE, Soto M, Ramírez R, Hernandez L: **A Type II Protein Secretory Pathway Required for Levansucrase Secretion by *Gluconacetobacter diazotrophicus*.** *J Bacteriol* 2004, **186**:5031-5039.
55. Menéndez C, Banguela A, Caballero-Mellado J, Hernández L: **Transcriptional Regulation and Signal-Peptide-Dependent Secretion of Exolevanase (LsdB) in the Endophyte *Gluconacetobacter diazotrophicus*.** *Appl Environ Microbiol* 2009, **75**:1782-1785.
56. Attwood MM, van Dijken JP, Pronk JT: **Glucose metabolism and gluconic acid production by *Acetobacter diazotrophicus*.** *J Ferment Bioeng* 1991, **72**:101-105.
57. Luna MF, Bernardelli CE, Galar ML, Boiardi JL: **Glucose metabolism in batch and continuous cultures of *Gluconacetobacter diazotrophicus* PAI 3.** *Cur Microbiol* 2006, **52**:163-168.
58. Luna MF, Mignone CF, Boiardi JL: **The carbon source influences the energetic efficiency of the respiratory chain of N2-fixing *Acetobacter diazotrophicus*.** *Appl Microbiol Biotechnol* 2000, **54**:564-569.
59. Matsushita K, Toyama H, Yamada M, Adachi O: **Quinoproteins: structure, function, and biotechnological applications.** *Appl Microbiol Biotechnol* 2002, **58**:13-22.
60. Miyazaki T, Sugisawa T, Hoshino T: **Pyrrroloquinoline quinone-dependent dehydrogenases from *Ketogulonicigenium vulgare* catalyze the direct conversion of L-sorbose to L-ascorbic acid.** *Appl Environ Microbiol* 2006, **72**:1487-1495.
61. Cascales E, Christie PJ: **The versatile bacterial type IV secretion systems.** *Nat Rev Microbiol* 2003, **1**:137-149.

62. Li PL, Hwang I, Miyagi H, True H, Farrand SK: **Essential components of the Ti plasmid *trb* system, and a type IV macromolecular transporter.** *J Bacteriol* 1999, **181**:5033-5041.
63. Thorsted PB, Macartney DP, Akhtar P, Haines AS, Ali N, Davidson P, Stafford T, Pocklington MJ, Pansegrau W, Wilkins BM, et al.: **Complete sequence of the IncP beta plasmid R751: implications for evolution and organization of the IncP backbone.** *J Mol Biol* 1998, **282**:969-990.
64. Mor M, Finger L, Stryker J, Fuqua C, Eberhard A, Winans S: **Enzymatic synthesis of a quorum sensing autoinducer through use of defined substrates.** *Science* 1996, **272**:1655-8.
65. Dougherty BA, Hill C, Weidman JF, Richardson DR, Venter JC, Ross RP: **Sequence and analysis of the 60 kb conjugative, and bacteriocin-producing plasmid pMRC01 from *Lactococcus lactis* DPC3147.** *Mol Microbiol* 1998, **29**:1029-1038.
66. Juhas M, Crook DW, Hood DW: **Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence.** *Cell Microbiol* 2008, **1012**:2377-2386.
67. Merritt PM, Danhorn T, Fuqua C: **Motility and chemotaxis in *Agrobacterium tumefaciens* surface attachment and biofilm formation.** *J Bacteriol* 2007, **189**:8005-8014.
68. Tomich M, Planet PJ, Figurski DH: **The *tad* locus: postcards from the widespread colonization island.** *Nat Rev Microbiol* 2007, **5**:363-375.
69. Rouws LF, Simões-Araújo JL, Hemerly AS, Baldani JL: **Validation of a Tn5 transposon mutagenesis system for *Gluconacetobacter diazotrophicus* through characterization of a flagellar mutant.** *Arch Microbiol* 2008, **189**:397-405.
70. Gillis M, Kersters K, Hoste B, Janssens D, Kroppenstedt RM, Stephan MP, Teixeira KRS, Döbereiner J: ***Acetobacter diazotrophicus* sp., a nitrogen-fixing acetic acid bacterium associated with sugarcane.** *Int J Systematic Bacteriol* 1989, **39**:361-364.
71. Rodrigues Neto J, Malavolta VA Jr, Victor O: **Meio simples para o isolamento e cultivo de *Xanthomonas campestris* pv. Citri Tipo B.** *Summa Phytopathologia* 1986, **12**:16.
72. Bonfield JK, Smith KF, Staden R: **A new DNA sequence assembly program.** *Nucleic Acids Res* 1995, **23**:4992-4999.
73. Loureiro MM, Bertalan M, Turque AS, Franca LM, Pádua VLM, Baldani JL, Martins OB, Ferreira PCG: **Physical and genetic map of the *Gluconacetobacter diazotrophicus* PAL5 chromosome.** *Rev Lat Microbiol* 2008, **50**:19-28.
74. Delcher AL, Bratke KA, Powers EC, Salzberg SL: **Identifying bacterial genes and endosymbiont DNA with Glimmer.** *Bioinformatics* 2007, **23**:673-679.
75. The UniProt Consortium: **The universal protein resource (UniProt).** *Nucleic Acids Res* 2008, **36**:190-195.
76. Hulo N, Bairoch A, Bulliard V, Cerutti L, Cuche B, De Castro E, Lachaise C, Langendijk-Genevaux PS, Sigrist CJA: **The 20 years of PROSITE.** *Nucleic Acids Res* 2007, **36**:D245-9.
77. Finn RD, Tate J, Mistry J, Coggill PC, Sammut JS, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A: **The Pfam protein families database.** *Nucleic Acids Res* 2008, **36**:281-288.
78. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, et al.: **InterPro: the integrative protein signature database.** *Nucleic Acids Res* 2009, **37**:211-215.
79. Emanuelsson O, Brunak S, von Heijne G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nat Protoc* 2007, **2**:953-971.
80. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305**:567-680.
81. Lowe TM, Eddy SR: **tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence.** *Nucleic Acids Res* 1997, **25**:955-964.
82. Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream MA: **Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database.** *Bioinformatics* 2008, **24**:2672-2676.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



Antioxidant pathways are up-regulated during biological nitrogen fixation to prevent ROS-induced nitrogenase inhibition in *Gluconacetobacter diazotrophicus*

Sylvia M. C. Alquéres · Jose Henrique M. Oliveira · Eduardo M. Nogueira · Helma V. Guedes · Pedro L. Oliveira · Fernando Câmara · Jose I. Baldani · Orlando B. Martins

Received: 25 March 2010 / Revised: 17 June 2010 / Accepted: 19 July 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract *Gluconacetobacter diazotrophicus*, an endophyte isolated from sugarcane, is a strict aerobe that fixates N₂. This process is catalyzed by nitrogenase and requires copious amounts of ATP. Nitrogenase activity is extremely sensitive to inhibition by oxygen and reactive oxygen species (ROS). However, the elevated oxidative metabolic rates required to sustain biological nitrogen fixation (BNF) may favor an increased production of ROS. Here, we explored this paradox and observed that ROS levels are, in fact, decreased in nitrogen-fixing cells due to the up-regulation of transcript levels of six ROS-detoxifying genes. A cluster analyses based on common expression patterns revealed the existence of a stable cluster with 99.8% similarity made up of the genes encoding the α -subunit of nitrogenase Mo-Fe protein (nitroD), superoxide dismutase (sodA) and catalase

type E (katE). Finally, nitrogenase activity was inhibited in a dose-dependent manner by paraquat, a redox cyler that increases cellular ROS levels. Our data revealed that ROS can strongly inhibit nitrogenase activity, and *G. diazotrophicus* alters its redox metabolism during BNF by increasing antioxidant transcript levels resulting in a lower ROS generation. We suggest that careful controlled ROS production during this critical phase is an adaptive mechanism to allow nitrogen fixation.

Keywords *Gluconacetobacter diazotrophicus* · Biological nitrogen fixation · Reactive oxygen species · Nitrogenase

Introduction

Sugarcane production in many countries depends on the traditional practice of adding nitrogen (NH₃) fertilizer for optimum growth and yield. In regions of the world where nitrogen input is very low, the diazotrophic bacterium *Gluconacetobacter diazotrophicus* has often been isolated from surface-sterilized roots, stems and leaves (Baldani et al. 1997; Cavalcante and Dobereiner 1988). Studies trying to understand the importance of *G. diazotrophicus* in plants grown under these conditions showed that in a N-deficient environment, sugarcane plants inoculated with the strain PA15 grew better and had a higher total N content than plants inoculated with the mutant nif- or uninoculated plants (Sevilla et al. 2001). These results indicate that the transfer of fixed N from *G. diazotrophicus* to sugarcane might be a significant mechanism to support plant growth.

BNF is the process by which N₂ is converted into NH₃ and is considered an energy-demanding process because it consumes 16 mol of ATP per mol of N₂ (Kim and Rees

Communicated by Stuart Ferguson.

S. M. C. Alquéres (✉) · J. H. M. Oliveira · P. L. Oliveira · O. B. Martins
Laboratório de Biologia Molecular,
Instituto de Bioquímica Médica,
Universidade Federal do Rio de Janeiro,
Rio de Janeiro, RJ CEP 21941-590, Brazil
e-mail: alqueres@bioqmed.ufrj.br

E. M. Nogueira
Depto de Genética e Biologia Molecular,
Universidade Federal do Estado do Rio de Janeiro,
Rio de Janeiro, RJ, Brazil

H. V. Guedes · J. I. Baldani
Laboratório de Genética e Bioquímica,
Embrapa Agrobiologia, Seropédica, RJ, Brazil

F. Câmara
Instituto Microbiologia Prof. Paulo de Góes.,
Universidade Federal do Rio de Janeiro,
Rio de Janeiro, RJ, Brazil

1994). This energy requirement can be met by aerobic respiration (Flores-Encarnacion et al. 1999). In this process, energy-rich substrates (nutrients) are oxidized, donating its electrons to a series of membrane proteins, known as the electron transport chain (ETC.), which will conserve the energy of the electron flow in the form of a membrane potential that will be the driving force to allow ATP synthesis. During respiration, molecular oxygen is the final electron acceptor, receiving 4 electrons, being reduced to water. However, some electrons may escape from the ETC., giving rise to partially reduced oxygen intermediates, commonly called reactive oxygen species (ROS) (Kowaltowski et al. 2009). Despite the need of high-energy supplies to support BNF, which is only achieved by respiration, nitrogenase is notoriously sensitive to oxygen and ROS (Robson and Postgate 1980).

So far, studies concerning the effects of oxygen on *G. diazotrophicus* have mainly dealt with the nature of its respiratory system. The most accepted hypothesis, called “respiratory protection”, postulates that the increase in O₂ consumption during BNF would allow sufficient ATP production and at the same time lower intracellular oxygen levels, preserving nitrogenase activity (Dalton and Postgate 1968; Flores-Encarnación et al. 1999; Pan and Vessey 2001; Ureta and Nordlund 2002). Less data have been acquired concerning the influence of increased respiration on the production and detoxification of reactive oxygen species (ROS), a natural by-product of aerobic metabolism and its effects on BNF. The generation of partially reduced forms of oxygen (i.e. ROS), such as superoxide radical (O₂⁻) and hydrogen peroxide (H₂O₂), may be part of the O₂ toxicity problem in nitrogen-fixing organisms (Dalton 1995). Considering the lack of information about free radical biology during BNF and the importance of this process for sustainable agriculture, the aim of our work is to investigate the influence of ROS production and detoxification in *G. diazotrophicus*. Our data revealed that, besides the respiratory protection mechanism, *G. diazotrophicus* alters its redox metabolism during BNF by increasing antioxidant transcript levels resulting in a decreased ROS generation that is critical to sustain nitrogenase activity and allow BNF. As we show that ROS can strongly inhibit the BNF process, we suggest that careful control of ROS production during this critical phase is an adaptive mechanism to allow nitrogen fixation.

Materials and methods

Cultivation conditions

Gluconacetobacter diazotrophicus strain PAL5 (BR11281) was obtained from the Embrapa Agro-biological Culture

Collection and was grown in LGI-P medium on a shaker with gentle agitation (100 rpm) for 48 h at 30°C. After three successive rounds of cultivation, cells were harvested by centrifugation at 12,000g for 2 min and resuspended in LGI-P medium (0.2 g/l K₂HPO₄, 0.6 g/l KH₂PO₄, 0.2 g/l MgSO₄ · 7H₂O, 0.02 g/l CaCl₂ · 2H₂O, 0.002 g/l Na₂MoO₄ · 2H₂O, 0.01 g/l FeCl₃ · 6H₂O, 100 g/l saccharose) with 1.0 mM of (NH₄)₂SO₄ to yield an A₆₀₀ of 0.6. Aliquots (1 ml) of cell suspensions were transferred into 250 ml flasks containing 50 ml of the fresh LGI-P medium (supplemented with 1.0 mM (nitrogen-fixing condition) or 20 mM of (NH₄)₂SO₄ (non-nitrogen-fixing condition) then incubated at 30°C with agitation (100 rpm).

The resistance of *G. diazotrophicus* cells to oxidative stress was investigated by the addition of Paraquat (PQ), a superoxide generator (Hassett et al. 1987) to a final concentration of 0.5 or 5 mM in fixing and non-nitrogen-fixing cultures after 24 h cultivation.

Determination of ROS production in *G. diazotrophicus*

Free-radical production was assessed by fluorescence microscopy of cells from liquid cultures grown in LGI-P medium under fixing and non-nitrogen-fixing conditions. *G. diazotrophicus* cultures were threefold concentrated and incubated for 10 min in the presence of the redox-sensitive dye CM-H₂DCFDA (5 μM) (Invitrogen—Molecular probes), followed by a washing step with fresh media and direct visualization (without fixative agents) of the samples. Images were acquired on a Zeiss Axioskop microscope with an Axiocam MRC5 using a Zeiss-09 filter set (excitation—BP 450–490; beam splitter—FT 510; emission—LP 515). Magnification was set at 630×.

Catalase activity

G. diazotrophicus was assayed for catalase activity according to (Aebi 1984). Briefly, crude extracts were incubated in Tris-buffer pH 7.0 in the presence of 9 mM hydrogen peroxide (H₂O₂) and absorbance was monitored at 240 nm for 1 min. The amount of protein was determined according to Bradford (1976).

Quantitative real-time polymerase chain reaction (qPCR)

The expression of selected genes was determined by real-time qPCR with SYBR Green Supermix (Applied Biosystems) using iCycler thermal cycler (Applied Biosystems). RNA was isolated with Trizol according to the manufacturer’s protocol. Samples were treated with DNase I (Invitrogen). Total RNA (2.5 μg) was used for cDNA synthesis with the Superscript First-strand synthesis kit

153 (Invitrogen). cDNA was purified using MinElute PCR
 154 purification kit (Qiagen) and used as qPCR templates. Each
 155 PCR reaction contained 12.5 μ l of 2 \times SYBR Green
 156 Supermix (Applied Biosystems), 1 μ M primers and
 157 appropriate templates in a 25- μ l reaction. PCR reactions
 158 were heated to 95°C for 3 min and then for 40 cycles with
 159 steps of 95°C for 30 s, 60°C for 15 s, 60°C for 45 s. The
 160 generation of specific PCR products was confirmed by the
 161 melting curve analysis and gel electrophoresis. The $2^{-\Delta\Delta C_t}$
 162 method (Livak and Schmittgen 2001) was employed for
 163 relative quantification and 23S rDNA was the housekeep-
 164 ing gene control. The results were based on the average of
 165 triplicate experiments. Primer sequences are available upon
 166 request.

167 Cluster analysis

168 We used qPCR to measure mRNA levels of cells grown
 169 under fixing or non-fixing conditions at 18, 24, 36, 48 and
 170 72 h. The expressions of *sodA*, *katE*, *kat*, *katC*, *gorA*, *gorB*
 171 and *nifD* were measured at the same time points and the
 172 cluster analysis was also carried out at these times. Genes
 173 were grouped on the basis of their common expression
 174 patterns across the time points and different conditions.

175 Cluster analysis of gene expression was carried out with
 176 the MINITAB v.14 package, using Euclidian distances and
 177 the Ward method for linkage. The single, complete, aver-
 178 age, centroid or median linkage methods displayed the
 179 same results in the test.

180 Nitrogenase activity

181 For the in vivo nitrogenase assay, 2 ml of *G. diazotro-*
 182 *phicus*-saturated culture grown in LGI-P medium was used
 183 to inoculate 250-ml flasks containing 48 ml LGI-P sup-
 184 plemented with 1 mM of (NH₄)₂SO₄. The cells were grown
 185 in a shaking water bath at 100 rpm for the indicated time at
 186 30°C. The nitrogenase activity was determined in whole
 187 cells by the acetylene reduction method as described (Reis
 188 and Dobereiner 1998). Ethylene produced was analyzed by
 189 gas chromatography in a Perkin Elmer gas chromatograph
 190 equipped with a flame ionization detector and Porapak N
 191 column. The protein content was determined by the
 192 Bradford method (Bradford 1976).

193 Results

194 ROS levels are decreased during nitrogen fixation

195 To test our hypothesis that redox metabolism is altered
 196 during the BNF process, we evaluated the presence of
 197 oxygen radicals in *G. diazotrophicus*. Cells were stimulated

to fix nitrogen and compared to a non-nitrogen-fixing 198
 condition. These two groups were loaded in the presence of 199
 the redox-sensitive fluorescent dye and observed under the 200
 microscope. Figure 1 shows that nitrogen-fixing cells 201
 (FIX—in Fig. 1) present a decreased fluorescent signal (1B 202
 and 1D), an indicative of lower ROS levels, when com- 203
 pared to non-nitrogen-fixing cells (NFIX) (1F and 1H), 204
 both at 24 and 48 h incubation time. 205

206 These results demonstrate that cellular ROS are reduced
 when *G. diazotrophicus* is fixing nitrogen, pointing to a 207
 possible inverse correlation between the presence of ROS 208
 and nitrogenase activity. Furthermore, it is evident the non- 209
 fixing cells inhabit an environment where ROS is part of 210
 the cellular milieu without exhibiting any noticeable 211
 adverse effect. This is consistent with the notion that ROS 212
 act as physiological molecules in a diverse set of biological 213
 situations (Forman et al. 2010; Terada 2006), as it seems to 214
 occur in non-fixing *G. diazotrophicus*. It is tempting to 215
 speculate why ROS levels have to be reduced during BNF 216
 and we hypothesized that the sensitivity of nitrogenase to 217
 inhibition by oxygen intermediates may be in the center of 218
 this question. 219

ROS detoxification is up-regulated during BNF process 220

221 In order to investigate the mechanisms involved in the
 decrease of ROS during BNF, we first measured catalase 222
 activity, an antioxidant enzyme involved in H₂O₂ removal. 223
 Figure 2a shows increased activity in nitrogen-fixing cul- 224
 tures at 48 and 72 h. Considering that catalase activity was 225
 increased up to sixfold in fixing cells and at the same time 226
 ROS levels were greatly reduced (Fig. 1), we hypothesized 227
 that other antioxidant pathways may also be active during 228
 BNF. Genes that could possibly be involved in ROS 229
 detoxification were identified in the recently published *G.* 230
diazotrophicus genome (Bertalan et al. 2009). Our analysis 231
 revealed the presence of six genes whose products are 232
 directly involved in ROS removal (Table 1). 233

234 qPCR used to measure mRNA levels of the identified
 ROS-detoxification genes in PAL-5 strain grown in high 235
 (NFIX) or low-nitrogen (FIX) media showed a strong 236
 induction of *sodA*, *katE*, *kat*, *katC* and *gorA* genes in 237
 response to nitrogen limitation (Fig. 2b). Importantly, 238
 genes that are not related to ROS metabolism (GS— 239
 glutamine synthetase—and amidase) do not have their 240
 expression levels altered during nitrogen fixation 241
 (Fig. 2b). Our results demonstrated a specific response of 242
 cells to reduce ROS levels under BNF conditions. Con- 243
 sistent with data presented in Fig. 1, non-fixing cultures 244
 seems to live in the presence of ROS and this condition 245
 does not represent a challenge since all the antioxidant 246
 enzymes analyzed were present in basal levels compared 247
 to fixing cells. 248

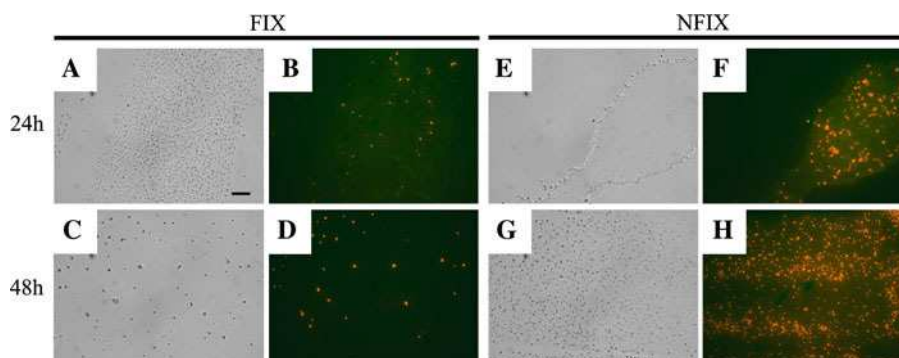


Fig. 1 ROS levels are decreased in nitrogen-fixing cells. *G. diazotrophicus* grown under fixing (*FIX*) and non-fixing (*NFIX*) conditions were stained with CM-H2DCFDA, sensitive to ROS. **a, c, e** and **g** represent *bright field* images. **b, d, f, h** are the *fluorescent* images, an

indicative of the presence of ROS. All images were acquired with the same exposure time to allow comparison of signal intensities. *Scale bar* 10 μ m

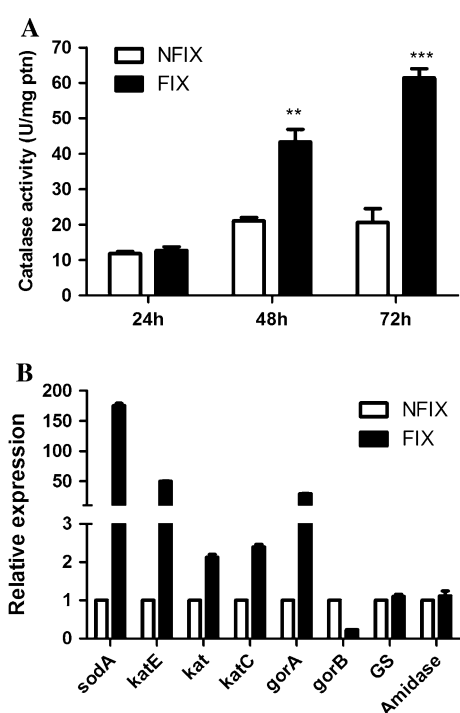


Fig. 2 Expression and activity of ROS detoxification in *G. diazotrophicus* during nitrogen fixation. **a** Catalase activity of cells grown in fixing and non-fixing conditions. **b** mRNA expression of the genes involved in ROS-detoxification from fixing and non-fixing cells. The data are expressed as the relative expression of the respective mRNAs normalized to the housekeeping gene 23S. The data are expressed as the average of three replicates \pm error bars

249 The genes for catalase type E (*katE*) and superoxide
250 dismutase A (*sodA*) are co-expressed with the α -subunit
251 of nitrogenase Mo-Fe protein (*nifD*)

252 In order to identify the most relevant ROS-detoxification
253 genes to BNF, we searched for distinct expression patterns
254 (clusters) (D'Haeseleer 2005). Genes were grouped on the
255 basis of their common expressions across time points and

under nitrogen-fixing and non-nitrogen-fixing conditions 256
(Fig. 3). Cluster analysis separated these genes in two 257
major clusters: one formed by *gorB* alone and the other 258
formed by *gorA*, *kat*, *katC*, *nifD*, *sodA* and *katE*, with 259
similarities up to 98% inside this group. Notice that the 260
genes present in this cluster are strongly induced as men- 261
tioned earlier (Fig. 2b). This last cluster can be divided into 262
three minor clusters *kat* + *gorA*; *katC*; and *katE* + 263
sodA + *nifD*. This cluster has 99.8% similarity and it is 264
stable when we use different algorithms for clustering 265
(single and average linkage, median linkage and ward 266
linkage), suggesting that *sodA* and *katE* have a major 267
protective role in the nitrogen fixation process. It is inter- 268
esting to note that the genes grouped in this cluster have 269
higher similarity to related enzymes from other non-sym- 270
biotic closely related subclass *Alphaproteobacteria* 271
(Table 1), while the other ROS-detoxifying genes analyzed 272
have higher similarity to related enzymes from phyloge- 273
netically distant symbiotic organisms. This could be an 274
indication that nitrogen fixation is an ancient process in 275
G. diazotrophicus and was probably acquired before the 276
adaptation to the endophytic lifestyle (Bertalan et al. 2009). 277

The presence of ROS strongly inhibits nitrogen fixation 278

To further explore differences concerning ROS levels 279
between nitrogen-fixing and non-fixing bacteria, the resis- 280
tance of *G. diazotrophicus* to oxidative stress was investi- 281
gated by the addition of Paraquat (PQ), a redox cyclor that 282
increases cellular ROS generation, into the cultures. The 283
growth of nitrogen-fixing cells (Fig. 4a) was severely 284
inhibited by the addition of PQ, indicating that the anti- 285
oxidant capacity of fixing cells (Fig. 2) is overwhelmed by 286
the pro-oxidant effects of this molecule. On the other hand, 287
non-nitrogen-fixing cells were resistant to both 0.5 and 288
5 mM of PQ, as they maintained high growth rates and 289
achieved high cell densities (Fig. 4a). General appearance 290

Table 1 Identification of antioxidant genes from *G. diazotrophicus*

Gene	Locus tag	Product	Conserved domains	Similar to
<i>sodA</i>	GDI2168	Superoxide dismutase	“Iron/manganese sod, C-terminal domain” “Iron/manganese sod, alpha-hairpin domain”	<i>Gluconobacter oxydans</i> (340s, 80%id)
<i>katE</i>	GDI0079	Catalase HPII	“KatE” “heme-binding pocket” “tetramer interface”	<i>Chromobacter salexigens</i> KatE (788s, 74%id)
<i>kat</i>	GDI0467	Non-heme catalase	“Catalase superfamily”	<i>Methylobacterium extorquens</i> Kat (332s, 53%id)
<i>katC</i>	GDI2359	Catalase HPII	“(GATase1)-like”	<i>Azoarcus</i> sp. KatC (1055s, 72%id)
<i>gorA</i>	GDI2216	Glutathione reductase	“Pyridine nucleotide-disulphide oxidoreductase” “Glutathione reductase”	<i>Acetobacter pasteurianus</i> (664s, 70%id)
<i>gorB</i>	GDI2280	Glutathione reductase	“Pyridine nucleotide-disulphide oxidoreductase” “HI0933-like protein”	<i>Xanthomonas campestris</i> (436s, 63%id)

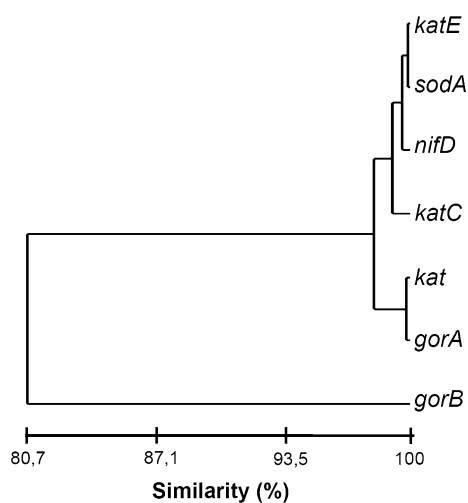


Fig. 3 Cluster analysis of redox genes. Notice that there were in two major clusters: one formed by *gorB* alone and the other formed by *gorA*, *kat*, *katC*, *nifD*, *sodA*, and *katE*, with similarities up to 98% inside this group. This last cluster can be divided in three minor clusters *kat* + *gorA*, *katC* and *katE* + *sodA* + *nifD*

291 of cells was similar for all PQ concentrations (data not
292 shown). This data is in accordance with previous results
293 (Figs. 1, 2) and demonstrate the resistance of non-fixing
294 cells to ROS also reveals that fixing cells up-regulate
295 antioxidant pathways (Fig. 2) to lower ROS levels (Fig. 1),
296 a condition that is essential for its survival (Fig. 4a).

297 To investigate whether BNF suppression was the basis
298 for inhibition of cell growth in the presence of PQ, nitro-
299 genase activity of cultures grown in the presence of PQ was
300 tested by the acetylene reduction essay. Cells grown
301 without PQ were used as a control. PQ was added to the
302 medium after 24 h of growth and the nitrogenase activity
303 was measured after a further 48 h (72 h of cultivation).
304 Nitrogenase activity decreased 64% when 0.5 mM of PQ
305 was added to the medium. When the concentration was
306 changed from 0.5 to 5 mM, nitrogenase activity decreased
307 to zero showing that the presence of ROS can strongly

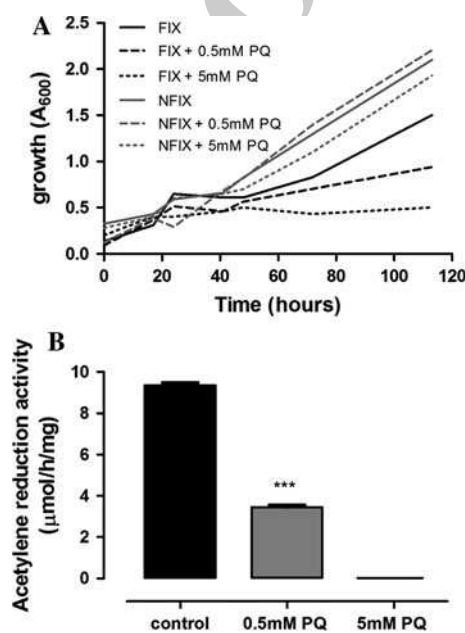


Fig. 4 ROS effect on cell growth and BNF activity. **a** Effect of different concentrations of PQ on the growth of fixing (black) and non-fixing (gray) cells. After an initial growth period of 24 h, PQ was added at a concentration of 0.5 mM or 5 mM (dashed and dotted lines, respectively). **b** Acetylene reduction activity (ARA) of *G. diazotrophicus* after 72 h cultivation. The activity was measured after 48 h in the presence of 0.5 and 5 mM of PQ

inhibit BNF in *G. diazotrophicus* (Fig. 4b). Therefore, 308
fixing cells up-regulate antioxidant pathways in order to 309
lower ROS levels and preserve nitrogenase activity, plac- 310
ing the regulation of ROS levels as a critical step to allow 311
biological nitrogen fixation. 312

Discussion 313

Here, we demonstrate that antioxidant pathways are 314
up-regulated during BNF to prevent ROS-induced 315

316 nitrogenase inhibition in *G. diazotrophicus*. Organisms that
 317 fix nitrogen usually adopt different strategies to protect
 318 nitrogenase from contact with O₂. The most studied is the
 319 *Rhizobia* strategy that infects the plant and forms nodules
 320 that create an anaerobic environment that favors BNF,
 321 while the plant supplies the energy requirement for the
 322 N-fixing bacteria (Fischer 1994). Differently from *Rhizobia*,
 323 *G. diazotrophicus* colonizes the intercellular spaces of the
 324 plant (apoplastic) and doesn't form nodules and is therefore
 325 always exposed to higher levels of O₂ (Dong et al. 1994).
 326 In fact, optimal nitrogen fixation by *G. diazotrophicus*
 327 demands high aerobic conditions both in solid (Pan and
 328 Vessey 2001) and liquid cultures (Flores-Encarnacion et al.
 329 1999). However, when oxygen is measured in the liquid
 330 phase during BNF, the maximum nitrogenase activity of
 331 *G. diazotrophicus* is observed at a pO₂ of 0.2 kPa (Reis and
 332 Dobreiner 1998). This is an indication that in order to
 333 keep the low levels of O₂ dissolved in the medium
 334 *G. diazotrophicus* adopts the so-called "respiratory pro-
 335 tection" strategy. Accordingly, *G. diazotrophicus* changes
 336 its electron transport chain composition during BNF. In this
 337 context, well-aerated cultures express cytochrome a1 and
 338 cytochrome bb as the main terminal oxidase. During
 339 repression of diazotrophic activity, cytochrome a1 dimin-
 340 ishes dramatically concomitantly with the appearance of
 341 cytochrome bd (Flores-Encarnacion et al. 1999). Oxidase
 342 activities are also much higher in membrane preparations
 343 obtained from cultures under BNF conditions than in those
 344 from cultures under non-nitrogen-fixing conditions and the
 345 respiratory rate was among the highest ever reported for
 346 aerobic bacteria (Flores-Encarnacion et al. 1999; Gonzalez
 347 et al. 2006).

348 However, an aspect that has been overlooked is the
 349 impact of the respiratory protection on ROS generation.
 350 Higher respiratory rates may lead to increased production
 351 of oxygen radicals (Boveris and Chance 1973), a condition
 352 that would favor redox imbalance and consequently oxi-
 353 dative stress. Here, we tried to explore this paradox and
 354 incorporate our data into the respiratory protection
 355 hypothesis. We propose that in parallel to the increase in
 356 aerobic respiration that occurs during BNF, *G. diazotro-*
 357 *phicus* also up-regulate antioxidant pathways to protect
 358 nitrogenase from ROS inhibition.

359 The importance of redox homeostasis to symbiotic BNF
 360 process has been indicated by previous studies with *Sino-*
 361 *rhizobium meliloti* strains. Mutants affected in the antiox-
 362 idant defense did not reach the differentiation stage of
 363 nitrogen-fixing bacteroids (Santos et al. 2000; Jamet et al.
 364 2003; Harrison et al. 2005). Similarly, a peroxiredoxin
 365 (prxS)/bifunctional catalase–peroxidase (*katG*) *Rhizobium*
 366 *etli* double mutant has a significantly reduced symbiotic
 367 nitrogen fixation capacity (Dombrecht et al. 2005; for a
 368 review see Chang et al. 2009). Still, these effects could be

369 both due to the presence of ROS or due to the incapacity of
 370 the bacteria to promote a normal nodule formation and
 371 therefore the environment that is necessary to allow
 372 nitrogen fixation. Our data show that ROS can directly
 373 inhibit nitrogenase activity, and *G. diazotrophicus* alters its
 374 redox metabolism during BNF to cope with that.

375 Concerning our gene analysis, most of the ORFs shared
 376 by *G. diazotrophicus* and other closely associated *Alpha-*
 377 *proteobacteria* are related to energy metabolism and are
 378 generally part of the core genome of *G. diazotrophicus*
 379 (Bertalan et al. 2009). One interesting aspect of catalase
 380 gene analysis (GDI0467 and GDI2359) is that although the
 381 most similar sequences are from phylogenetically distant
 382 organisms, they were all isolated from plant leaves and
 383 have the ability to promote the growth of various plant
 384 seedlings (Abanda-Nkpwatt et al. 2006; Sessitsch et al.
 385 2005). This is an indicative that these genes were acquired
 386 later, possible by lateral gene transfer. This could be
 387 especially important for bacterial adaptation to the endo-
 388 phytic lifestyle and may confer advantages to *G. diazo-*
 389 *trophicus* in comparison with other microbes that colonize
 390 the same niche.

391 This study can be regarded as the first step toward
 392 understanding of ROS metabolism in *Acetobacteraceae*. It
 393 is also the first characterization of ROS production and
 394 modification of the redox state during BNF process. Our
 395 main finding is that during BNF *G. diazotrophicus* activate
 396 antioxidant pathways that diminish the intracellular levels
 397 of ROS protecting nitrogenase from ROS toxicity. These
 398 results complement the respiratory protection hypothesis
 399 previously described and allow a better comprehension of
 400 nitrogenase activity in aerobic environments.

401 **Acknowledgments** This work was supported by CNPq, FINEP and
 402 FAPERJ.

403 **Open Access** This article is distributed under the terms of the
 404 Creative Commons Attribution Noncommercial License which per-
 405 mits any noncommercial use, distribution, and reproduction in any
 406 medium, provided the original author(s) and source are credited.
 407

408 References

- 409 Abanda-Nkpwatt D, Musch M, Tschiersch J, Boettner M, Schwab W
 410 (2006) Molecular interaction between *Methylobacterium extor-*
 411 *quens* and seedlings: growth promotion, methanol consumption,
 412 and localization of the methanol emission site. *J Exp Bot*
 413 57:4025–4032
 414 Aebi H (1984) Catalase in vitro. *Methods Enzymol* 105:121–126
 415 Baldani JI, Caruso L, Baldani VLD, Goi SR, Dobreiner J (1997)
 416 Recent advances in BNF with non-legume plants. *Soil Biol*
 417 *Biochem* 29:911–922
 418 Bertalan M, Albano R, de Padua V, Rouws L, Rojas C, Hemerly A,
 419 Teixeira K, Schwab S, Araujo J, Oliveira A et al (2009)
 420 Complete genome sequence of the sugarcane nitrogen-fixing

- 421 endophyte *Gluconacetobacter diazotrophicus* Pal5. BMC
422 Genomics 10:450
- 423 Boveris A, Chance B (1973) The mitochondrial generation of
424 hydrogen peroxide. General properties and effect of hyperbaric
425 oxygen. Biochem J 134:707–716
- 426 Bradford MM (1976) A rapid and sensitive method for the
427 quantitation of microgram quantities of protein utilizing the
428 principle of protein-dye binding. Anal Biochem 72:248–254
- 429 Cavalcante VA, Dobreiner J (1988) A new acid-tolerant nitrogen-
430 fixing bacterium associated with sugarcane. Plant Soil
431 108:23–31
- 432 Chang C, Damiani I, Puppo A, Frendo P (2009) Redox changes
433 during the legume–*Rhizobium* symbiosis. Mol Plant 2:370–377
- 434 D’Haeseleer P (2005) How does gene expression clustering work?
435 Nat Biotechnol 23:1499–1501
- 436 Dalton DA (1995) Antioxidant defenses of plants and fungi. Chapman
437 & Hall, New York
- 438 Dalton H, Postgate JR (1968) Effect of oxygen on growth of
439 *Azotobacter chroococcum* in batch and continuous cultures.
440 J Gen Microbiol 54:463–473
- 441 Dombrecht B, Heusdens C, Beullens S, Verreth C, Mulkers E, Proost
442 P, Vanderleyden J, Michiels J (2005) Defence of *Rhizobium etli*
443 bacteroids against oxidative stress involves a complexly regu-
444 lated atypical 2-Cys peroxiredoxin. Mol Microbiol 55:1207–1221
- 445 Dong Z, Canny MJ, McCully ME, Roboredo MR, Cabadilla CF,
446 Ortega E, Rodes R (1994) A nitrogen-fixing endophyte of
447 sugarcane stems (a new role for the apoplast). Plant Physiol
448 105:1139–1147
- 449 Fischer HM (1994) Genetic regulation of nitrogen fixation in rhizobia.
450 Microbiol Rev 58:352–386
- 451 Flores-Encarnacion M, Contreras-Zentella M, Soto-Urzua L, Aguilar
452 GR, Baca BE, Escamilla JE (1999) The respiratory system and
453 diazotrophic activity of *Acetobacter diazotrophicus* PAL5.
454 J Bacteriol 181:6987–6995
- 455 Forman HJ, Maiorino M, Ursini F (2010) Signaling functions of
456 reactive oxygen species. Biochemistry 49:835–842
- 457 Gonzalez B, Martinez S, Chavez JL, Lee S, Castro NA, Dominguez
458 MA, Gomez S, Contreras ML, Kennedy C, Escamilla JE (2006)
459 Respiratory system of *Gluconacetobacter diazotrophicus* Pal5.
460 Evidence for a cyanide-sensitive cytochrome bb and cyanide-
461 resistant cytochrome ba quinol oxidases. Biochim Biophys Acta
462 1757:1614–1622
- 463 Harrison J, Jamet A, Muglia CI, Van de Sype G, Aguilar OM, Puppo
464 A, Frendo P (2005) Glutathione plays a fundamental role in
growth and symbiotic capacity of *Sinorhizobium meliloti*.
J Bacteriol 187:168–174
- Hasset DJ, Britigan BE, Svendsen T, Rosen GM, Cohen MS (1987)
Bacteria form intracellular free radicals in response to paraquat
and streptonigrin. Demonstration of the potency of hydroxyl
radical. J Biol Chem 262:13404–13408
- Jamet A, Sigaud S, Van de Sype G, Puppo A, Herouart D (2003)
Expression of the bacterial catalase genes during *Sinorhizobium*
meliloti–*Medicago sativa* symbiosis and their crucial role during
the infection process. Mol Plant Microbe Interact 16:217–225
- Kim J, Rees DC (1994) Nitrogenase and biological nitrogen fixation.
Biochemistry 33:389–397
- Kowaltowski AJ, de Souza-Pinto NC, Castilho RF, Vercesi AE
(2009) Mitochondria and reactive oxygen species. Free Radic
Biol Med 47:333–343
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression
data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method.
Methods 25:402–408
- Pan B, Vessey JK (2001) Response of the endophytic diazotroph
Gluconacetobacter diazotrophicus on solid media to changes in
atmospheric partial O₂ pressure. Appl Environ Microbiol
67:4694–4700
- Reis VM, Dobreiner J (1998) Effect of high sugar concentration on
nitrogenase activity of *Acetobacter diazotrophicus*. Arch Micro-
biol 171:13–18
- Robson RL, Postgate JR (1980) Oxygen and hydrogen in biological
nitrogen fixation. Annu Rev Microbiol 34:183–207
- Santos R, Hérouart D, Puppo A, Touati D (2000) Critical protective
role of bacterial superoxide dismutase in *Rhizobium*–legume
symbiosis. Mol Microbiol 38:750–759
- Sessitsch A, Coenye T, Sturz AV, Vandamme P, Barka EA, Salles JF,
Van Elsland JD, Faure D, Reiter B, Glick BR et al (2005)
Burkholderia phytofirmans sp. nov., a novel plant-associated
bacterium with plant-beneficial properties. Int J Syst Evol
Microbiol 55:1187–1192
- Sevilla M, Burris RH, Gunapala N, Kennedy C (2001) Comparison of
benefit to sugarcane plant growth and ¹⁵N₂ incorporation
following inoculation of sterile plants with *Acetobacter diazo-*
trophicus wild-type and Nif-mutants strains. Mol Plant Microbe
Interact 14:358–366
- Terada LS (2006) Specificity in reactive oxidant signaling: think
globally, act locally. J Cell Biol 174:615–623

SYLVIA MARIA CAMPBELL ALQUÉRES

Rua Prudente de Moraes 564/305 – Rio de Janeiro, RJ, Brasil. CEP: 22420-040
alqueres@bioqmed.ufrj.br, salqueres@gmail.com

FORMAÇÃO ACADÊMICA

2006 – 2010	Doutorado em Química Biológica - Universidade Federal do Rio de Janeiro.
2006 – 2008	Especialização em Filosofia Contemporânea - Pontifícia Universidade Católica do Rio de Janeiro.
2004 – 2006	Mestrado em Química Biológica - Universidade Federal do Rio de Janeiro.
1998 - 2003	Graduação em Microbiologia e Imunologia - Universidade Federal do Rio de Janeiro.

IDIOMAS

Português (língua materna), Inglês (fluido), Alemão (bom), Francês (bom), Espanhol (bom).

BOLSAS DE ESTUDO

2006 - 2010:	Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq (doutorado)
07/2008 – 10/2008:	NASA Planetary Biology Internship CNPq
2004 – 2006:	Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq (mestrado)
1999 – 2002:	Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq (iniciação científica)

EXPERIÊNCIA EM PESQUISA

2006 – Present	Universidade Federal do Rio de Janeiro – Laboratório de Biologia Molecular, Instituto de Bioquímica Médica, Rio de Janeiro, RJ, Brasil. Genômica funcional de <i>Gluconacetobacter diazotrophicus</i> . Orientador: Dr. Orlando B. Martins.
07/2008 – 10/2008	NASA Ames Research Center –Protein Nanotechnology Group, Moffett Field, CA, EUA. Estudo de função e estabilidade da proteína HSP60 de archaeas. Orientador: Dr. Jonathan D. Trent.
10/2004 – 03/2005	Universidade Federal do Rio de Janeiro – Laboratório de Biologia Molecular, Instituto de Bioquímica Médica, Rio de Janeiro, RJ, Brasil.

Clonagem e expressão de genes de archaeas. Caracterização de enzimas de archaeas. Síntese de proteínas em extremófilos.

Orientador: Dr. Orlando B. Martins.

1999 – 2000

Universidade Federal do Rio de Janeiro – Laboratório de Tecnologia Enzimática, Instituto de Química, Rio de Janeiro, RJ, Brasil.

Desenvolvimento de meios não convencionais para o crescimento de archaeas hipertermofílicas. Produção de enzimas degradadoras de amido por archaeas.

Orientador: Dr. Carolina M. M. C. Andrade.

PUBLICAÇÕES

ALQUÉRES, SMC; OLIVEIRA, JHM; NOGUEIRA, EN; GUEDES, HV; OLIVEIRA, PL; CÂMARA, FC; BALDANI, JI; MARTINS, OB (2010) Antioxidant pathways are up-regulated during biological nitrogen fixation to prevent ROS-induced nitrogenase inhibition in *Gluconacetobacter diazotrophicus*. *Archives of Microbiology*.

FERREIRA, PC *et al.* (2009) Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* PAL5. *BMC Genomics* 10: 450.

ALQUÉRES, SMC; BRANCO, RV; FREIRE, DMG; ALMEIDA, RV; MARTINS, OB. Characterization of a thermostable esterase (Pf2001) from *Pyrococcus furiosus* expressed in *Escherichia coli*: effects of thioredoxin fusion tag and Triton X-100. *Submitted*.

ALMEIDA, RV; BRANCO, RV; PEIXOTO, B; LIMA, C; **ALQUÉRES, SMC**; ANTUNES, O; MARTINS, OB; FREIRE, DMG (2007) Immobilization of a thermostable esterase (Pf2001) from *Pyrococcus furiosus* on microporous polypropylene: isotherms, hyperactivation and purification. *Biochemical Engineering Journal* 39: 531-537.

ALQUÉRES, SMC; ALMEIDA, RV; CLEMENTINO, MBM; VIEIRA, RP; ALMEIDA, WI; CARDOSO, AM; MARTINS, OB (2007) Exploring the biotechnological applications in the archaeal domain. *Brazilian Journal of Microbiology* 38: 398-405.

ALMEIDA, RV; **ALQUÉRES, SMC**; LARENTIS, AL; RÖSSLE, SCS.; CARDOSO, AM; ALMEIDA, WI; BISCH, PM; ALVES, TLM; MARTINS, OB (2006) Cloning, expression, partial characterization and structural modeling of a novel esterase from *Pyrococcus furiosus*. *Enzyme and Microbial Technology* 39:1128-1136.

CARDOSO, AM; CLEMENTINO, MM; MARTINS, OB; VIEIRA, RP; ALMEIDA, RV; **ALQUÉRES, SMC**; ALMEIDA, WI (2003) Archaea: Potencial Biotecnológico. *Biotecnologia ciência & desenvolvimento*. 30:71-77.

PRODUTOS TECNOLÓGICOS

ALMEIDA, RV; FREIRE, DMG; MARTINS, OB; ALVES, TLM; **ALQUÉRES, SMC**; BRANCO, RV (2008) PI-0800268-1 (registro sob sigilo).

TEXTOS EM JORNAIS DE NOTÍCIA/ REVISTAS

ALQUÉRES, SMC; MARTINS, OB (2008) A nova empreitada de Venter: rumo à vida artificial. *Ciência Hoje*, v. 248, Março.

ALQUÉRES, SMC; ALMEIDA, RV (2005) O feijão e a vida no limite. *Ciência hoje das crianças*, v. 155.

OUTRAS ATIVIDADES

Produção do livro *Uma breve história das vacinas - ensinando ciência com arte* dos autores Vivian Rumjaneck and artist Aurélio Corrêa (Rio de Janeiro: Imos, 2008).

Organização da **Semana de Pós-graduação em Bioquímica I & II**. Rio de Janeiro, 2007 & 2008.

Presidente da **Sociedade Brasileira de Bioquímica e Biologia Molecular (SBBq) Jovem** 2007-2008.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)