



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS DA SAÚDE
INSTITUTO DE BIOFÍSICA CARLOS CHAGAS FILHO

Letícia Miranda Lery Santos

Análise proteômica da bactéria *Gluconacetobacter diazotrophicus*: estabelecimento de um banco proteômico, análise diferencial e implicações na anotação do genoma

**Rio de Janeiro
Fevereiro de 2007**

Livros Grátis

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

Milhares de livros grátis para download.

Letícia Miranda Lery Santos

Análise proteômica da bactéria *Gluconacetobacter diazotrophicus*: estabelecimento de um banco proteômico, análise diferencial e implicações na anotação do genoma

Dissertação de mestrado submetida ao Instituto de Biofísica Carlos Chagas Filho da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de mestre em Ciências Biológicas – Biofísica

Orientador: **Paulo Mascarello Bisch**
Co-orientadora: **Wanda Maria Almeida von Krüger**

Unidade Multidisciplinar de Genômica - Instituto de Biofísica
Carlos Chagas Filho - Universidade Federal do Rio de Janeiro

Lery, LMS

Análise proteômica da bactéria *Gluconacetobacter diazotrophicus*: estabelecimento de um banco proteômico, análise diferencial e implicações na anotação do genoma / Leticia Miranda Lery Santos. Rio de Janeiro: UFRJ, IBCCF, 2007.

V, 183f., il.;

Orientador: Paulo Mascarello Bisch; Co-orientadora: Wanda Maria Almeida von Krüger

Dissertação de Mestrado – UFRJ / IBCCF / Ciências Biológicas – Biofísica, 2007.

Referências Bibliográficas: 129-144.

1. *Gluconacetobacter diazotrophicus*. 2. Proteoma.

I. Bisch, Paulo Mascarello. II. Universidade Federal do Rio de Janeiro, Instituto de Ciências Biomédicas. III. Análise proteômica da bactéria *Gluconacetobacter diazotrophicus*: estabelecimento de um banco proteômico, análise diferencial e implicações na anotação do genoma.

RESUMO

As tecnologias de proteômica e bioinformática, associadas às ferramentas de anotação de genomas, foram utilizadas para descrever principalmente o potencial metabólico da bactéria endofítica *Gluconacetobacter diazotrophicus*. Esta bactéria, encontrada no interior de plantas como cana-de-açúcar, café, abacaxi e batata-doce, fixa o nitrogênio atmosférico, produz hormônios de crescimento vegetal e bacteriocinas, sendo também capaz de solubilizar compostos de zinco, contribuindo para o crescimento acentuado de seus hospedeiros. Nesta dissertação foi estabelecido o primeiro banco proteômico de *G. diazotrophicus*. Com as técnicas de eletroforese uni e bidimensional, cromatografia líquida e espectrometria de massas, foram identificadas 573 proteínas e 40 vias metabólicas ativas. Proteínas envolvidas em processos de adaptação celular, respostas de proteção e mecanismos regulatórios também foram descritas. Utilizando técnicas de proteômica diferencial foram também analisadas 31 proteínas especificamente expressas na fase exponencial ou estacionária do crescimento de *G. diazotrophicus* e 26 com expressão dependente da disponibilidade de compostos nitrogenados. Na fase exponencial de crescimento, além do metabolismo de DNA e processos de produção e conversão de energia, foram identificadas proteínas envolvidas na homeostasia do pH intracelular. Por outro lado, na fase estacionária, as células responderam ao estresse da limitação de nutrientes, pela redução das atividades anabólicas e parada dos sistemas de geração de energia. Já em limitação de compostos nitrogenados, foram encontradas proteínas envolvidas na fixação de nitrogênio, tais como NifB, NifW e NifR3, e outras envolvidas na habilidade simbiótica de *G. diazotrophicus*. Quando em abundância de compostos nitrogenados, *G. diazotrophicus* expressou uma proteína do sistema urease, indicando uma estratégia adicional para captação desses compostos. Por fim, os métodos de separação de proteínas e a espectrometria de massas foram utilizados como ferramentas para o aprimoramento da anotação do genoma de *G. diazotrophicus*, possibilitando a validação experimental das 573 ORFs que codificam as proteínas identificadas, a correção de 84 N-terminais e 25 C-terminais de proteínas pelo seqüenciamento de peptídeos semi-trípticos, detecção de 5 erros de *frameshift*, de 3 ORFs melhor previstas numa versão anterior do banco do genoma, e o assinalamento de 13 ORFs não previstas pelo programa Glimmer.

ABSTRACT

Proteomics, bioinformatics and genomic techniques were used to investigate metabolic pathways of the endophytic bacterium *Gluconacetobacter diazotrophicus*. This bacterium, found within plants such as sugarcane, pineapple, coffee and sweet potato, fixes atmospheric nitrogen, produces plant growth-promoting hormones and bacteriocins and helps solubilization of zinc compounds. The first proteomic view of *G. diazotrophicus* PAL5 is presented herein, showing 573 proteins and 40 active metabolic pathways. Most of these proteins are involved in intermediary metabolism, although several proteins implicated in cellular adaptation, protection responses and regulators were also found. Moreover, proteins specifically induced at exponential and stationary phases of cultures, under low and high levels of nitrogen compounds, were described. During exponential growth phase, under both conditions, *G. diazotrophicus* differentially expressed genes involved in energy and DNA metabolism and also proteins involved in the cytoplasmic pH homeostasis. Most abundant proteins in stationary cells were transporters, transferases and stress associated. Under low levels of nitrogen compounds, cells expressed many nitrogen-fixation related proteins, such as NifB, NifW and NifR3, and some typical proteins of symbiotic relationships, such as NodC and NodM. Interestingly, two hydrogenases were detected, suggesting that nitrogen fixation is optimized in *G. diazotrophicus* by H₂ recycling. Among proteins differentially induced under high level of combined nitrogen we detected UreD, an accessory factor for the assembly of urease, an enzyme which catalyzes the hydrolysis of urea to ammonia and carbamate. Urea based fertilizers are largely used in the agricultural practices, therefore, UreD expression might suggest that *G. diazotrophicus* has the potential to uptake ammonium from other sources in the environment. Additionally, mass spectrometry data allowed the experimental validation of 573 ORFs, assignment of 13 new ORFs, correction of 5 frameshifts and the adjustment of 86 N or C-terminal of ORFS in the *G. diazotrophicus* genome.

AGRADECIMENTOS

Agradeço a minha família por ter me proporcionado essa vida tão bela e todas as oportunidades de alcançar meus objetivos. Ao meu namorado, Paulo Ricardo, pelo exemplo que é e pela alegria de todos os dias. Aos amigos dos vários laboratórios, por tornar o dia a dia cada vez mais curioso. Aos meu orientadores, Paulo e Wanda, e a nossa colaboradora, Ana, por todo o suporte e discussão científica.

LISTA DE SIGLAS

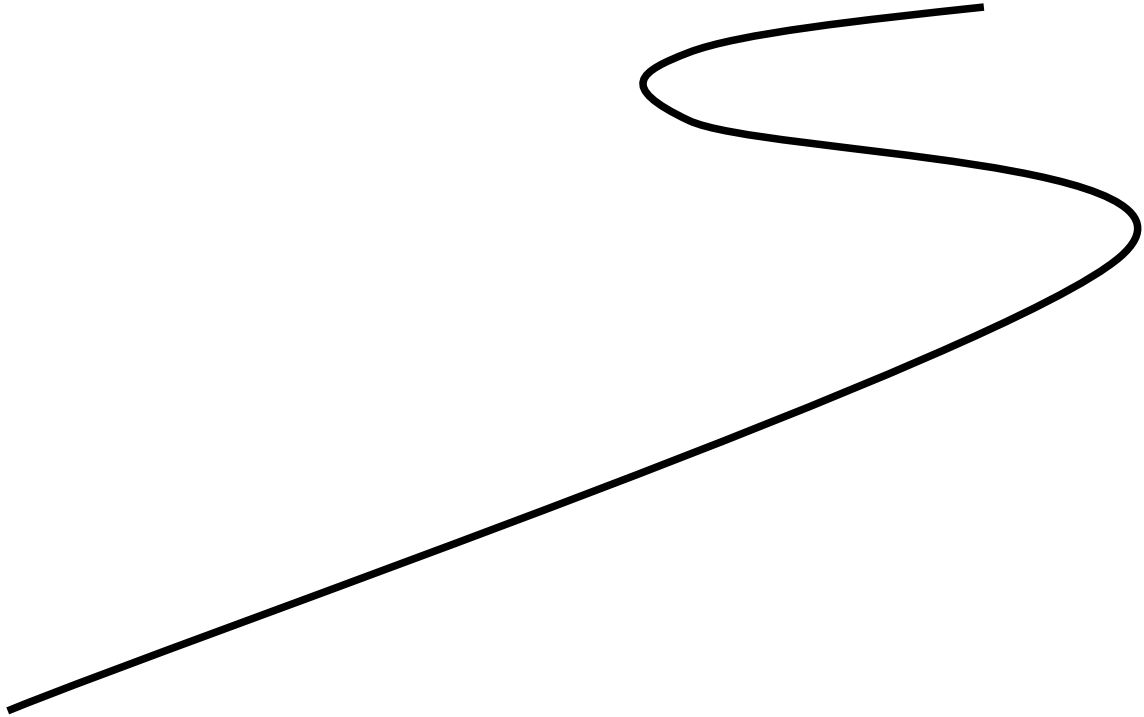
2D-GE	Eletroforese bidimensional em gel de poliacrilamida
ACN	Acetonitrila
CHAPS	3-[(3-Colamidopropil)dimetilamonio]-1-propanosulfonato
DTT	Ditiotreitol
DYGS	Meio de cultura rico para o isolamento de <i>Xanthomonas campestris</i>
EC	<i>Enzyme comission number</i> (Kegg)
ESI	Ionização por spray de elétrons (<i>Electronspray ionization</i>)
FA	Ácido fórmico
LC	Cromatografia líquida (<i>Liquid Cromatography</i>)
LGI	Meio de cultura Lipman e Baldani
LGI-P	Meio de cultura Lipman e Baldani - Pernambuco
MALDI	Ionização por desorção a laser assistida por uma matriz (<i>Matrix Assisted Laser Desorption Ionization</i>)
MM	Massa molecular
MS	Espectrometria de massas
MS/MS	Espectrometria de massas em seqüência
MudPit	Multidimensional protein identification technology
ORF	Fase aberta de leitura (<i>Open reading frame</i>)
pI	Ponto isoelétrico
PMSF	Fluoreto de metil fenil sulfonato
SDS	Sódio dodecil sulfato
TFA	Ácido trifluoroacético
TOF	Tempo de vôo (<i>Time of flight</i>)

ÍNDICE

1- INTRODUÇÃO.....	12
1.1- <i>Gluconacetobacter diazotrophicus</i>	12
1.1.1- A descoberta, hospedeiros e dispersão.....	12
1.1.2- Classificação e características gerais.....	14
1.1.3- Caracterização metabólica.....	15
1.1.4- Fixação de nitrogênio.....	17
1.1.5- Produção de bacteriocina.....	23
1.1.6- Produção de hormônios de crescimento vegetal.....	25
1.1.7- Solubilização de compostos de zinco.....	26
1.1.8- Importância de <i>G. diazotrophicus</i> para o desenvolvimento sustentável.....	27
1.2- Projetos Genoma.....	30
1.2.1- Estado da arte.....	30
1.2.2- O projeto Riogene.....	31
1.3- Projetos proteoma.....	32
1.3.1- Estado da arte.....	32
1.3.2- Rede Proteômica do Rio de Janeiro (PROTEOMA-RIO).....	38
2- OBJETIVOS.....	41
3- MATERIAIS E MÉTODOS.....	43
3.1- Cepa e condições de cultivo.....	43
3.2- Lise celular.....	44
3.3- Separação de proteínas.....	45
3.4- Análise de imagens.....	46
3.5- Espectrometria de massas.....	47
3.6- Identificação de proteínas.....	49
3.7- Análise bioinformática.....	50
4 – RESULTADOS E DISCUSSÃO.....	52
4.1- BANCO PROTEÔMICO.....	54
4.1.1 – Crescimento celular e obtenção das amostras.....	54
4.1.2- Eficiência de detergentes na lise celular.....	54
4.1.3- Construção do banco de proteínas de <i>G. diazotrophicus</i>	55
4.1.3.1- Eletroforese bidimensional e espectrometria de massas MALDI.....	55
4.1.3.2- Análise de MudPit.....	58
4.1.4– Localização e categorização funcional das proteínas identificadas.....	63
4.1.4.1- Proteínas e membrana.....	63
4.1.4.2– Categorização das proteínas.....	63
4.1.4.3– Vias metabólicas ativas de <i>G. diazotrophicus</i>	65
4.1.4.4- Proteínas envolvidas em transporte e outras proteínas de superfície celular..	71
4.1.4.5- Proteínas relacionadas à fixação de nitrogênio.....	72
4.1.4.6- Proteínas de adaptação e resposta de proteção.....	72
4.1.4.7- Proteínas reguladoras.....	74
4.1.5– Conclusões parciais.....	74
4.2– ANÁLISE DIFERENCIAL.....	77
4.2.1– Cinética de crescimento.....	77

4.2.2– Análise proteômica: comparação de perfis protéicos obtidos por eletroforese bidimensional	79
4.2.3- Identificação das proteínas expressas diferencialmente.....	85
4.2.4 – Categorização funcional das proteínas	86
4.2.5- Análise das proteínas induzidas na fase exponencial (E).....	91
4.2.6- Análise das proteínas induzidas na fase estacionária (S)	96
4.2.7- Análise das proteínas induzidas em limitação de compostos nitrogenados (L)...	100
4.2.8- Análise das proteínas induzidas em abundância de compostos nitrogenados (H)	103
4.2.9- Conclusões parciais	104
4.3- GENOMA X PROTEOMA	107
4.3.1 – Contribuições do proteoma para a anotação do genoma	107
4.3.2 – Conclusões parciais	122
5- CONCLUSÕES	125
REFERÊNCIAS BIBLIOGRÁFICAS	129
ANEXO.....	145
TABELA DO BANCO PROTEÔMICO.....	146

I - INTRODUÇÃO



1.1- A *Gluconacetobacter diazotrophicus*

1.2- Projetos Genoma

1.3- Projetos Proteoma

1- INTRODUÇÃO

1.1- A *Gluconacetobacter diazotrophicus*

1.1.1- A descoberta, hospedeiros e dispersão

G. diazotrophicus foi isolada pela primeira vez por Cavalcante e Döbereiner, em 1988 [1], a partir de raízes e caules de diferentes variedades de cana-de-açúcar, oriundas de diversas regiões do Brasil. A bactéria foi descrita como “uma nova bactéria fixadora de nitrogênio, tolerante ao ácido”, que inicialmente recebeu o nome de *Saccharobacter nitrocaptans* [1]. Ainda durante o período de impressão do artigo original e com base em experimentos de hibridização RNA/DNA e DNA/DNA, um novo nome foi proposto para esta bactéria: *Acetobacter nitrocaptans*. Em pouco tempo a própria Dr^a Johanna Döbereiner propôs a alteração do nome para *Acetobacter diazotrophicus* e em alguns meses o novo nome foi oficializado. Posteriormente, Yamada e colaboradores [2] propuseram um novo esquema de taxonomia da família Acetobacteriaceae, transferindo a espécie *Acetobacter diazotrophicus* para o gênero *Gluconacetobacter*. Desde então, o nome oficial é *Gluconacetobacter diazotrophicus*.

O isolamento inicial de *G. diazotrophicus* de raízes lavadas e do interior de talos de cana-de-açúcar [1] foi posteriormente confirmado pelo seu isolamento de tecidos internos de diferentes variedades de cana-de-açúcar cultivadas na Austrália e no México [3, 4]. Ainda, *G. diazotrophicus* foi isolada do inseto *Saccharococcus sacchari* [5], praga comum nas regiões onde se cultiva cana-de-açúcar, e em plantas de batata-doce (*Ipomoea batatas*) e grama *Pennisetum purpureum* var Cameroon [6]. A ausência de *G. diazotrophicus* na rizosfera de cana-de-açúcar assim como na rizosfera e interior de plantas cultivadas alternadamente com a cana-de-açúcar [7], sugeriram seu caráter endofítico e a idéia da associação específica com plantas que acumulam sacarose e se propagam de forma vegetativa [6]. O recente isolamento de *G. diazotrophicus* de plantas de abacaxi (*Ananas comosus*) suportam esta hipótese [8].

Recentemente, *G. diazotrophicus* tem sido detectada ocasionalmente na rizosfera de cana-de-açúcar [4] e também na rizosfera e tecidos internos de outras plantas como o café (*Coffea arabica*) [9], arroz [10] e do cereal *Eleusine coracana* [11], que aparentemente não são ricas em sacarose. O isolamento de *G. diazotrophicus* de plantas de famílias tão diversas como Poaceae, Convolvulaceae, Rubiaceae e Bromeliaceae sugerem uma ampla capacidade de dispersão e de colonização de hospedeiros por esta bactéria (tab. 1). Em 2004, Youssef e colaboradores [12] inocularam plantas de trigo com *Gluconacetobacter spp.* e observaram por microscopia sua dispersão nas raízes e caules, demonstrando a habilidade desta bactéria em colonizar um hospedeiro até então não identificado como natural [12].

Considerando que a cana-de-açúcar é originária da Nova Guiné, o café do Oriente Médio e o abacaxi da Amazônia, torna-se difícil estabelecer a origem da interação *G. diazotrophicus*-plantas hospedeiras. Tendo em vista que a espécie *G. diazotrophicus* é subdividida em 11 grupos eletroforéticos e que a diversidade populacional de *G. diazotrophicus* em café é maior do que a observada em outras plantas, sugere-se que a associação com café seja mais antiga do que as demais descritas [8]. Ainda, existe a possibilidade de que o hospedeiro primário de *G. diazotrophicus* não seja uma planta de interesse econômico e portanto não tenha sido identificado.

Tabela 1: Hospedeiros naturais de *G. diazotrophicus* [1, 4-11, 13-16]

<i>Hospedeiro</i>	<i>Região</i>	<i>Referências</i>
Cana-de-açúcar	Raízes, caules, folhas	1, 4, 7, 14
Batata doce	Raízes, caules	13
Café	Raízes, caules, rizosfera	9
Abacaxi	Fruto	8
Arroz	Raízes	10
Cochonilhas	Interno	5, 16
Chá	Raízes	15
Gramma	Raízes, caules	6
Ragi	Raízes, caules, rizosfera	11

Devido ao caráter endofítico de *G. diazotrophicus* em plantas que se reproduzem assexuadamente, propôs-se que sua ampla dispersão geográfica foi decorrente da reprodução vegetativa de plantas que naturalmente abrigavam a bactéria [17]. Ainda, sugere-se que a dispersão regional desta bactéria entre variedades de cana-de-açúcar tenha sido mediada por insetos como o *S. sacchari*, conhecido como cochonilha [5]. Vale ressaltar que a frequência de isolamento de *G. diazotrophicus* é maior na região apical de talos, justamente o local de alimentação das cochonilhas. Por fim, sugere-se que alguns fungos micorrízicos vesículo-arbusculares podem contribuir para transmissão e dispersão da *G. diazotrophicus*.

1.1.2- Classificação e características gerais

A *Gluconacetobacter diazotrophicus* é uma bactéria Gram-negativa e aeróbica que pertence à sub-classe das α -proteobactérias [1, 2]. Membro do gênero *Gluconacetobacter* da família *Acetobacteraceae* [2], essa espécie é capaz de oxidar etanol à ácido acético e apresenta-se em forma de bastonetes retos com extremidades arredondadas e dimensões

aproximadas de $0,7 \times 0,9 \times 2,0 \mu\text{m}$ [1, 15]. Os representantes da espécie *G. diazotrophicus* são móveis, contendo flagelos laterais ou peritríqueos (fig. 1) e podem ser distinguidos das outras espécies desse gênero por formarem colônias que apresentam pigmentação marrom com bordas claras em placas de meio de cultura Batata P (contendo 5 a 15% de sacarose) e alaranjada no meio definido LGIP contendo azul de bromotimol [18].

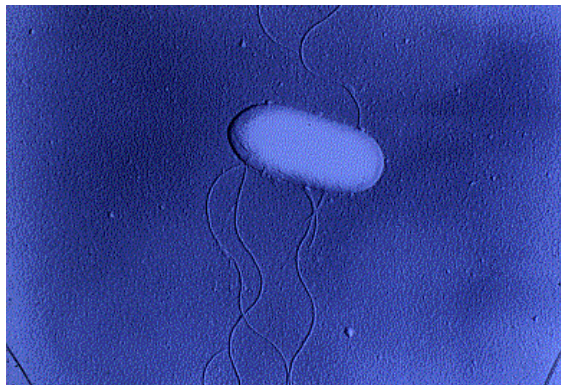


Figura 1: Microscopia eletrônica de transmissão, mostrando a bactéria endofítica *G. diazotrophicus* em sua forma de bastonete e seus flagelos (fonte: <http://www.cnpab.embrapa.br/servicos/baby/acetober.gif>).

O crescimento ótimo de *G. diazotrophicus* ocorre em meios que contêm 10% de sacarose, à temperatura de 30°C e o pH inicial de 5,5 [1]. Contudo, esta bactéria é capaz de crescer em concentrações de sacarose de até 30% e fixar nitrogênio em pH menores do que 3,0 [1]. Além disso, *G. diazotrophicus* pode utilizar glicose, frutose ou sacarose como fonte de carbono, produzindo ácidos orgânicos a partir desses substratos [1, 19]. Com exceção da sacarose, esta espécie não utiliza outros dissacarídeos e não é capaz de crescer com ácidos dicarboxílicos como única fonte de carbono [1].

1.1.3- Caracterização metabólica

Uma caracterização metabólica inicial da *G. diazotrophicus* foi realizada por Alvarez e Martínez-Drets [20]. Ensaio enzimáticos, a partir de lisados celulares da cepa PAL3, revelaram que a principal via metabólica ativa do metabolismo de carboidratos é a das pentoses fosfato [20]. Enzimas da via alternativa Entner Doudoroff não foram detectadas. Por

outro lado, as enzimas necessárias a fosforilação de hexoses e interconversões de hexoses fosforiladas, assim como frutose bifosfatase e outras enzimas da via gliconeogênica foram detectadas, sugerindo que glicose, frutose, glicerol e gluconato podem ser utilizados como fonte de carbono [20].

Além disso, *G. diazotrophicus* oxida diretamente glicose, frutose, gluconato e cetogluconatos através de enzimas da cadeia respiratória, sendo o O₂ aceptor final de elétrons [20]. A oxidação direta de glicose pode fazer parte de um mecanismo de proteção respiratória da nitrogenase ao oxigênio, visto que Stephan e colaboradores [21] demonstraram a tolerância da nitrogenase ao oxigênio em presença de glicose e gluconato. Contribuindo para este mesmo mecanismo, estariam formas da glicose desidrogenase, uma solúvel e outra ligada à membrana, sendo esta última responsável pela acidificação do meio de cultura [20]. A oxidação da glicose no espaço periplasmático e seu metabolismo também já foram descritas [22].

G. diazotrophicus possui ainda o ciclo do ácido tricarboxílico completo, mas a atividade da via do glioxilato não foi detectada [20]. Enzimas das vias de metabolismo de dicarboxilatos foram encontradas, embora *G. diazotrophicus* não seja capaz de crescer com dicarboxilatos, tais como succinato, fumarato e malato, como única fonte de carbono. Tal fato foi explicado pela ausência de um sistema de transporte de dicarboxilatos [20].

Em relação ao metabolismo da sacarose, não foi observado nenhum sistema de transporte pela membrana deste dissacarídeo. Contudo, uma enzima com atividade sacarolítica, a levansucrase, foi detectada em sobrenadante de cultura de *G. diazotrophicus* [23]. A especificidade e eficiência, assim como a caracterização molecular e estrutural desta enzima foram descritas em trabalhos posteriores [23-28].

1.1.4- Fixação de nitrogênio

Duas características peculiares desta espécie, que permitiram a sua caracterização microbiológica [1] são seu hábitat endofítico e sua capacidade de fixar o nitrogênio atmosférico. Organismos endofíticos são aqueles que vivem no interior de plantas sem lhes causar danos, apresentando localização intracelular ou intercelular, colonizando principalmente raízes, colmos e folhas ou o tecido vascular do hospedeiro [29]. Este modo de associação apresenta vantagens competitivas em relação as bactérias rizosféricas (vivem na superfície das raízes), considerando que a disponibilidade de nutrientes é maior no interior das plantas e o número de espécies endofíticas é menor do que o das rizosféricas [30]. Além disto, as bactérias endofíticas encontram-se mais protegidas das condições adversas do meio ambiente e, por estabelecerem um contato mais íntimo com seus hospedeiros, podem usufruir de benefícios mais diretos em comparação com bactérias rizosféricas [29, 30].

Geralmente, *G. diazotrophicus* é encontrada no apoplasto intercelular dos caules [17] e nos vasos do xilema da base do colmo de cana-de-açúcar [29], mas a ocorrência intracelular da bactéria em células de determinadas regiões das raízes da cana-de-açúcar também foi observada [31]. Os mecanismos pelos quais a bactéria entra na cana-de-açúcar ainda não foram conclusivamente estabelecidos [15]. *G. diazotrophicus* apresenta baixa sobrevivência em solos onde a cana-de-açúcar tem sido cultivada, embora tenha sido encontrada na rizosfera da cana-de-açúcar em solos australianos [4]. Segundo James e colaboradores [32] *G. diazotrophicus* inicialmente coloniza a superfície das raízes consumindo exudatos e posteriormente entra no apoplasto via junções laterais das raízes.

A fixação de nitrogênio por *G. diazotrophicus* tem sido assunto de diversos estudos [21, 33-36]. O nitrogênio é um dos principais componentes de biomoléculas, fazendo parte da estrutura de aminoácidos, bases nitrogenadas, hormônios e clorofila, entre outras moléculas, sendo portanto, essencial para a sobrevivência e crescimento dos organismos. O nitrogênio

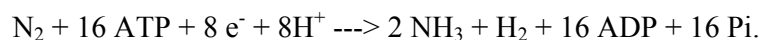
gasoso, N_2 , embora constitua quase 80% da atmosfera terrestre, é quimicamente inerte à temperatura ambiente e, diferentemente de outros elementos que ocorrem na natureza, suas reservas minerais são relativamente raras e a reserva de nitrogênio presente na matéria orgânica do solo é limitada, podendo ser esgotada rapidamente por cultivos sucessivos [37]. Além disso, em regiões tropicais, as condições de temperatura e umidade dos solos aceleram os processos de decomposição da matéria orgânica e perdas de nitrogênio, resultando em solos com baixos teores desse nutriente, entre 0,05 e 0,30% [38].

O processo de fixação de nitrogênio constitui-se na redução do nitrogênio atmosférico (N_2) a amônia (NH_3) e requer uma energia de ativação alta. Apesar de termodinamicamente favorável, a reação não ocorre espontaneamente sem a presença de catalisadores adequados, visto que os dois átomos de nitrogênio da molécula de N_2 encontram-se unidos através de uma tripla ligação, de uma maneira muito estável. Assim, processos naturais contribuem com apenas 10% do nitrogênio fixado, em grande parte decorrente de atividade vulcânica, combustão industrial e queima de florestas [39].

Por outro lado, Haber-Bosch desenvolveu um processo industrial para a síntese de amônia, que emprega altas temperatura (400 a 600°C) e pressão (100 a 200 atm) e utiliza catalisadores a base de ferro [37]. O gasto de fontes energéticas não renováveis na fixação industrial de nitrogênio é calculado em seis barris de petróleo por tonelada de NH_3 sintetizada [37].

Em contraposição, a fixação biológica do N_2 é mediada por uma ampla gama de microrganismos procarióticos, incluindo cianobactérias e bactérias Gram-positivas e Gram-negativas. Este processo ocorre graças a um complexo enzimático, denominado nitrogenase, presente nesses organismos [37]. Do ponto de vista energético, ele também é dispendioso para o organismo que o realiza. No entanto, devido à presença desse sistema enzimático, a reação

pode ocorrer à temperatura ambiente e pressão atmosférica. A estequiometria da reação proposta para redução de N_2 a duas moléculas NH_3 é:



(onde e^- simboliza elétron e Pi simboliza o fosfato inorgânico).

Visto que os organismos eucariontes são incapazes de absorver o N_2 e convertê-lo a uma forma assimilável, a maioria das plantas obtém o nitrogênio do solo sob a forma de íon nitrato (NO_3^-), havendo algumas que o absorvem sob a forma de íon amônio (NH_4^+). Porém, estas duas formas são extremamente instáveis no solo, representando o principal agravante da utilização de fertilizantes nitrogenados, ou seja uma baixa eficiência de aproveitamento pelas plantas, raramente ultrapassando 50% [40]. Isto significa um aumento do custo da produção pois metade dos fertilizantes nitrogenados aplicados à agricultura é perdida por lixiviação dos solos (lavagem no perfil do solo), denitrificação (transformação do NO_3^- em N_2 e NO_2^-) e pela volatilização do NH_3 .

Portanto, enquanto os processos naturais contribuem com 10% do nitrogênio fixado e a fixação industrial com cerca de 15% do nitrogênio atmosférico aplicado no solo, é o processo de fixação biológica do nitrogênio (FBN) que se encarrega de garantir a maior parte dos 75% restantes [39], sendo de suma importância para a nutrição nitrogenada das plantas e fertilidade dos solos.

A forma mais comum e extensivamente estudada da nitrogenase é um complexo dependente de molibdênio, consistindo de duas metalo-proteínas, a ferro-proteína (Fe-proteína ou dinitrogenase redutase) e a ferro-molibdênio-proteína (FeMo-proteína ou dinitrogenase). A Fe-proteína é a componente doadora de elétrons e ligadora de nucleotídeos, enquanto a FeMo-proteína contém o sítio redutor do substrato [41]. Há outras nitrogenases homólogas, no entanto, com o molibdênio substituído por vanádio ou ferro [42]. Um outro tipo de

nitrogenase, dependente de superóxido, foi isolado de *Streptomyces thermoautotrophicus* e é portanto diferente das demais classes de nitrogenases descritas [43].

A Fe-proteína do complexo nitrogenase Mo-dependente, codificada pelo gene *nifH* [41] é um homodímero que coordena um grupo 4Fe-4S. Depois da redução e ligação ao ATP, a Fe-proteína reduz a FeMo-proteína com concomitante hidrólise do ATP. A FeMo-proteína, codificada pelo genes *nifD* e *nifK* [41], é um tetrâmero com duas subunidades alfa (NifD) e duas subunidades beta (NifK), coordenadoras de dois cofatores FeMo contendo molibdênio-ferro-enxofre-homocitrato (FeMoco) e dois *clusters* 8Fe-7S. O *cluster* 8Fe-7S (P-*cluster*) é localizado na interface entre as subunidades NifD e NifK (fig. 2) e acredita-se que está envolvido no transporte de elétrons entre a Fe-proteína e o FeMoco na FeMo-proteína [44].

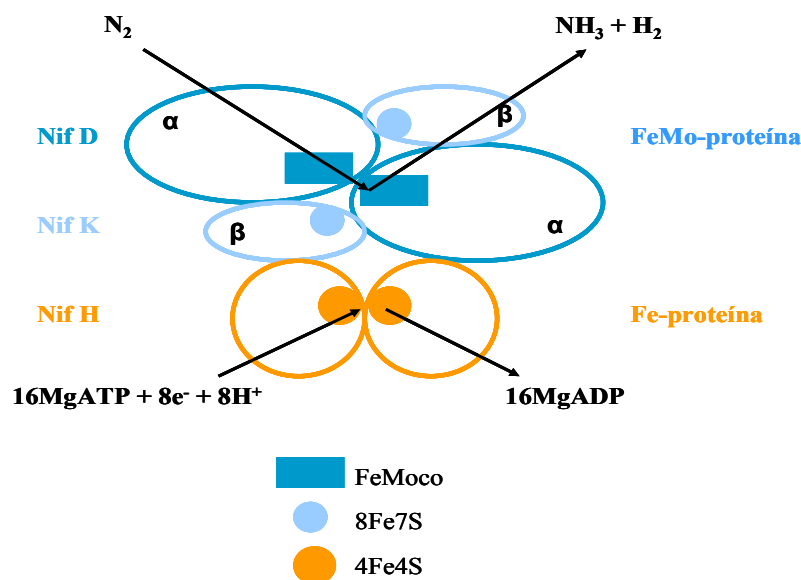


Figura 2: Esquema representativo do complexo nitrogenase. Em azul são mostradas as subunidades e cofatores da FeMo-proteína. Em laranja, o homodímero e o cofator da Fe-proteína.

Antes de uma molécula de nitrogênio ser reduzida a duas moléculas de amônia e, como subproduto, uma molécula de hidrogênio H_2 ser gerada, ocorrem pelo menos oito ciclos de redução. A partir dos dados obtidos com sistemas de nitrogenase em *Clostridium pasteurianum*, *Klebsiella pneumoniae* e *Azotobacter vinelandii*, descreveu-se uma seqüência geral de eventos representando o ciclo catalítico da nitrogenase [45]. Duas moléculas Mg-

ATP ligam a Fe-proteína reduzida, diminuindo o potencial de redução-oxidação do *cluster* Fe-S da Fe-proteína. A ligação de Mg-ATP também induz uma mudança conformacional que favorece a interação entre Fe-proteína e cada uma das subunidades da FeMo-proteína. Após a formação desse complexo, as duas moléculas de Mg-ATP são hidrolisadas a Mg-ADP, concomitante com a transferência de um elétron da Fe-proteína para a FeMo-proteína. [46-48]. Seguindo a transferência de elétrons, o complexo nitrogenase se dissocia, sendo este um passo limitante do ciclo. O ciclo catalítico é repetido até que um número suficiente de elétrons tenha sido transferido para reduzir o N₂ ligado ao FeMoco.

A nitrogenase pode ainda reduzir substratos alternativos, que tenham duplas ou triplas ligações em sua estrutura, como a molécula do N₂. Um desses substratos é o acetileno, reduzido a etileno. Uma vez que o envelope bacteriano é permeável tanto ao acetileno quanto ao etileno e o etileno pode ser facilmente medido por cromatografia gasosa, essa reação é amplamente utilizada para a determinação da atividade nitrogenase em extratos de componentes purificados, células inteiras e em plantas [49].

Os genes que codificam as subunidades da nitrogenase dependente de Mo foram identificados em *G. diazotrophicus* [34, 50, 51]. Assim como em outros diazotrofos eles fazem parte do operon *nif*HDK. Este operon faz parte de um grupo (*nif-fix*) de genes diretamente relacionados ou associados com o processo de fixação de nitrogênio, compreendidos em 30.5 kb do genoma de *G. diazotrophicus*, o maior grupo contíguo de genes *nif* caracterizado em bactérias diazotróficas [50]. Tais genes estão organizados em 8 unidades transcricionais [50] e o arranjo geral desses genes é semelhante ao encontrado em *Azospirillum brasiliense*. Entretanto, individualmente os produtos gênicos são mais similares aos das espécies de *Rhizobiaceae* ou de *Rhodobacter capsulatus*.

Além dos genes *nif*HDK, o grupo de genes *nif-fix* em *G. diazotrophicus* inclui: *nifA*, cujo produto é o ativador transcricional de *nif*HDK; *nifB*, *nifE*, *nifN*, *nifQ*, *nifV* e *nifX*, todos

envolvidos na síntese do cofator FeMo; *nifZ* e *nifS*, atuantes na ativação e maturação da nitrogenase; *nifU*, relacionado com ativação e maturação da nitrogenase e montagem dos *clusters* de Fe; *nifW*, envolvido com a proteção da Mo-Fe proteína ao oxigênio; *rpoN*, o fator σ^{54} ; *fixABCX*, codificadores de uma cadeia de transporte de elétrons [52]; *fdxN*, *fdxB*, *fixU*, *ynf1*, *orf1*, *orf4* e *orf6*, de função desconhecida, mas homólogos a genes de outros diazotrofos; *mcpA*, que codifica uma proteína que responde a sinais extracelulares e quimiotaxia [53]; *modABCD*, que codifica um sistema de transporte de molibdato em *E. coli* [54]; e *orf11*, que não apresenta similaridade com outros genes [50].

Quanto à tolerância a compostos nitrogenados em geral, a nitrogenase de *G. diazotrophicus* é parcialmente inibida por 20 mM de sulfato amônio, $(\text{NH}_4)_2\text{SO}_4$, especialmente por não possuir a enzima nitrato redutase [55] e é tolerante a vários aminoácidos [21]. A fixação de nitrogênio por *G. diazotrophicus* em meio líquido requer uma dose inicial de 0,6 a 1mM de $(\text{NH}_4)_2\text{SO}_4$, assim como a oxidação de glicose extracelular seguida da formação de ácido glucônico, essenciais ao início da fase logarítmica.

Outro fator envolvido na regulação da atividade da nitrogenase é a concentração de oxigênio. Ambas Fe-proteína e MoFe-proteína são danificadas irreversivelmente por altas concentrações de oxigênio [56]. Assim como outros diazotrofos aeróbicos, *G. diazotrophicus* desenvolveu diversas estratégias para evitar essa inibição irreversível. Flores-Encarnacion e colaboradores [57] observaram que durante a fixação de nitrogênio, as células de *G. diazotrophicus* aumentam a taxa respiratória, continuamente removendo o O_2 dissolvido nas culturas aeróbicas, provendo um ambiente intracelular adequado para fixação de N_2 . Em 2002, Dong [58] sugeriu que a estrutura de uma colônia e a localização da bactéria dentro da colônia teriam funções na proteção da nitrogenase ao fluxo excessivo de O_2 [58].

Em 2001, Pan e Vessey [59], estudaram os efeitos da pressão atmosférica de O_2 na atividade da nitrogenase de *G. diazotrophicus* e observaram que ela diminui rapidamente

quando a pressão de O₂ aumenta, mas é rapidamente recuperada com a diminuição da pressão de O₂, indicando que *G. diazotrophicus* apresenta um mecanismo reversível de proteção em resposta a variações da pO₂ [59]. Ureta e Nordlund [60], utilizando anticorpos contra a proteína FeSII de *A. vinelandii*, identificaram uma proteína de aproximadamente 14 kDa em extratos de células de *G. diazotrophicus* crescidas em condições de fixação de N aeróbicas. No entanto, em condições anaeróbicas, a proteína provável FeSII não foi detectada. Estes resultados foram considerados uma forte evidência de um mecanismo de proteção conformacional envolvendo uma provável proteína FeSII, similar ao mecanismo descrito em *A. vinelandii* [61]. Outros mecanismos de regulação da fixação de nitrogênio em *G. diazotrophicus* já foram descritos [16, 35, 58, 59, 62] e alguns serão discutidos adiante neste trabalho.

1.1.5- Produção de bacteriocina

Diversas metodologias, tais como o balanço de nitrogênio, diluição de ¹⁵N e a abundância natural de ¹⁵N, comprovaram os altos índices de fixação biológica de nitrogênio em plantas gramíneas, tais como cana-de-açúcar, arroz e grama forrageira. Alguns dos efeitos da inoculação de *G. diazotrophicus* em plantas de cana-de-açúcar também já foram observados (fig. 3), e incluem o aumento acentuado do crescimento radicular e da superfície foliar [63, 64]. Tais efeitos são dependentes da cepa de *G. diazotrophicus*, assim como da variedade de cana-de-açúcar [63, 64], sendo a cepa PAL5 de *G. diazotrophicus* a mais frequentemente isolada de cana-de-açúcar [64].

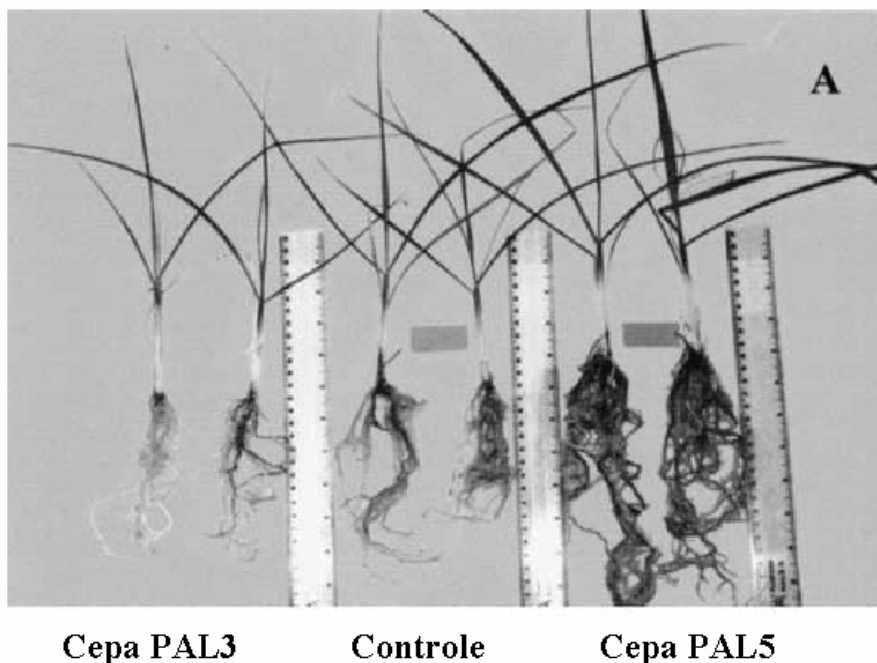


Figura 3: Efeito da inoculação das cepas PAL3 e PAL5 de *G. diazotrophicus* em plantas de cana-de-açúcar da variedade MEX 57-473. Comparação entre plantas inoculadas com as cepas PAL3 e PAL5 e controle, mostrando que a cepa PAL5 estimula o crescimento desta variedade de cana-de-açúcar, enquanto que a cepa PAL3 não causou alterações significativas [64].

No entanto, tais efeitos resultam de estratégias utilizadas pela *G. diazotrophicus* vão além do processo de fixação de nitrogênio. Em 2002, Piñon e colaboradores [65] identificaram uma proteína produzida e secretada pela *G. diazotrophicus* em cultura líquida, capaz de impedir o crescimento de *X. albilineans* (agente causador de queimaduras de folhas, uma doença bacteriana vascular da cana-de-açúcar [66]) em meio sólido (fig. 4). É sugerido que o modo de ação desta proteína ocorre pela destruição da camada fina mais interna da parede celular e pela indução de perda de conteúdo citoplasmático para o meio. A purificação dessa molécula levou a identificação de uma proteína “bacteriocina tipo lisozima” com massa em torno de 12,5 kDa, ligeiramente menor do que outras lisozimas cuja massa molecular varia em torno de 14 kDa.

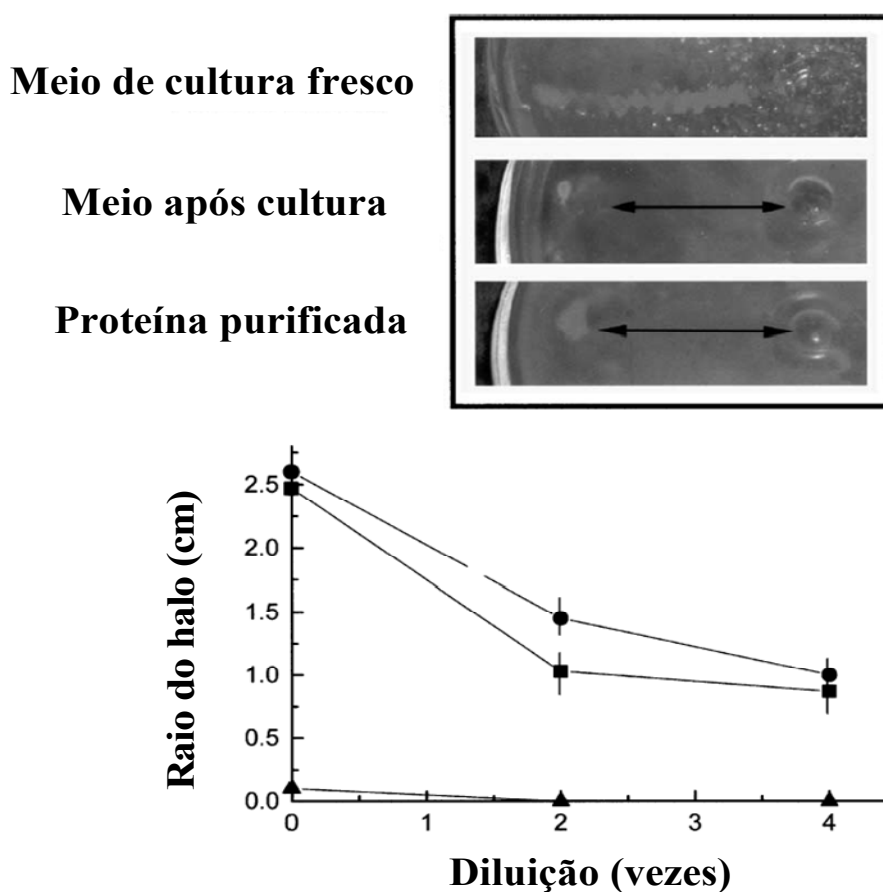


Figura 4: Halo de inibição do crescimento de *X. albilineans* por substância produzida e secretada no meio de cultura (■) ou da proteína isolada (●) de *G. diazotrophicus* [65]. O meio de cultivo fresco, ou seja, não inoculado, foi utilizado como controle negativo (▲).

1.1.6- Produção de hormônios de crescimento vegetal

Um terceiro fator de contribuição ao crescimento acentuado das plantas inoculadas com *G. diazotrophicus* pode ser a produção de fitohormônios pela bactéria. Em 1993, Fuentes-Ramirez e colaboradores [3] descreveram a produção de ácido-3-indol-acético, um hormônio que estimula o crescimento vegetal, por cepas de *G. diazotrophicus* isoladas de cultivares de cana-de-açúcar do México [3]. Bastian e colaboradores caracterizaram fitohormônios produzidos *in vitro* por esta bactéria, confirmando a produção de ácido-3-indol-acético (IAA) e indicando a produção de giberelinas GA1 e GA3 por *G. diazotrophicus* crescida em meio definido contendo 10-15% de sacarose [67].

1.1.7- Solubilização de compostos de zinco

A característica de *G. diazotrophicus* descrita mais recentemente, que também deve contribuir para a promoção do crescimento de plantas é a solubilização de compostos de zinco. O zinco é um micronutriente essencial para o crescimento de plantas, sendo cofator de diversas enzimas metabólicas. Seu suprimento em grande parte é devido ao uso de fertilizantes agrícolas, embora no solo o zinco rapidamente seja transformado em formas insolúveis. Assim, a solubilização de compostos de zinco por microorganismos é fundamental para manutenção deste hábitat saudável. As cepas PAL5 e L3 de *G. diazotrophicus* solubilizam com eficiência compostos insolúveis de zinco (fig. 5) [68]. Tal eficiência de solubilização é dependente da fonte de carbono, sendo mais pronunciada a 10% de glicose. Nessa condição, foi observada a presença de ácido 5-cetoglucônico no meio extracelular, que deve ser um mediador da solubilização do zinco [68]. Esta habilidade de *G. diazotrophicus* traz vantagens ecológicas importantes, primeiro por tornar o cátion disponível para as plantas em solos deficientes deste metal, e ainda pela possibilidade de usar esta bactéria em processos de fitoextração e fitorremediação de solos contaminados por zinco [68].

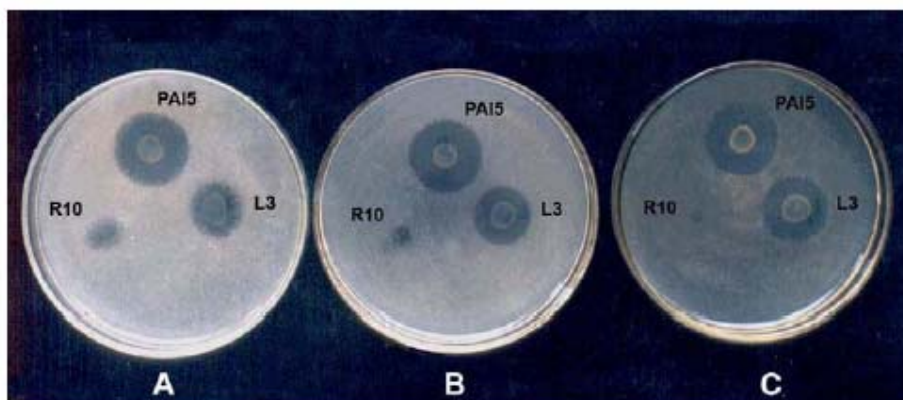


Figura 5: Halo de solubilização de compostos de zinco produzido por cepas de *G. diazotrophicus* (R10, PAL5 e L3) em meio LGI contendo 10 g de glicose por litro, suplementado com 0.1% (w/v) das fontes de zinco: (A) ZnO, (B) ZnCO₃, and (C) Zn₃(PO₄)₂ [68].

1.1.8- Importância de *G. diazotrophicus* para o desenvolvimento sustentável

A agroindústria é um dos principais segmentos da economia brasileira, com importância tanto no abastecimento interno como no desempenho exportador do Brasil. Uma avaliação recente estimou que sua participação no Produto Interno Bruto (PIB) representa cerca de 10%, considerando-se apenas o valor da produção e ocupa uma posição de destaque entre os diversos setores da economia. Quando se usa o conceito moderno de agronegócio, que abrange a soma das operações de produção e distribuição de insumos e novas tecnologias agrícolas, produção propriamente dita, armazenamento, transporte, processamento e distribuição dos produtos agrícolas e seus derivados, a participação do complexo agroindustrial alcança mais de 35% do PIB, evidenciando o efeito multiplicador que esse setor exerce sobre a economia, principalmente no interior do país [69]. No ano de 2003, a balança comercial do agronegócio registrou saldo recorde de US\$ 21,55 bilhões, 26,6% superior em relação ao saldo de 2002. As receitas de exportação cresceram 22,9% no mesmo período, alcançando a cifra de US\$ 25,52 bilhões, enquanto o volume exportado aumentou 13,5% [70].

Os sistemas de produção agrícola no Brasil, assim como a maioria dos sistemas econômicos embasados no desenvolvimento capitalista, são caracterizados pela maximização da produção por unidade de área cultivada (produtividade). Este paradigma, aliado aos avanços tecnológicos do século XX, levou ao que se costuma chamar de "Revolução Verde". Práticas agrícolas de baixa produtividade foram substituídas nos anos 1960/70 por sistemas de alta produtividade porém extremamente dependentes de insumos químicos, notadamente pesticidas e fertilizantes. A modernização da agricultura impulsionou o desenvolvimento de nações eminentemente agrícolas como o Brasil, capitalizando e modernizando a zona rural e aumentando a receita proveniente das exportações. Todavia, a preocupação meramente econômica destes modelos de produção desconsiderou os efeitos e conseqüências dessas

novas práticas agrícolas sobre o ambiente natural, causando grandes impactos no equilíbrio global a nível de diversos ecossistemas terrestres.

Ainda hoje, muitos economistas avaliam o desempenho do sistema agrícola brasileiro por meio de indicadores quantitativos parciais, tal como, quantidade de “insumos e tecnologias agrícolas” (por exemplo, pesticidas e fertilizantes) aplicados por área versus unidade de tempo. Esta visão do processo agrícola é equivocada e, aliada à eficiente tática de *marketing* das indústrias de pesticidas, em sua maioria multinacionais até recentemente descompromissadas com a conservação de nossos recursos naturais, levou o Brasil à condição de quinto maior aplicador de pesticidas do mundo [71]. Como consequência deste modelo de desenvolvimento agrícola podemos citar a contaminação e degradação de solos e águas, desertificação, salinização, redução da biodiversidade, desequilíbrios ecológicos e a crescente resistência de pragas e ervas daninhas aos pesticidas [71]. Essa resistência, associada à dependência de insumos químicos por parte dos produtores tem impulsionado a indústria à descoberta e formulação de novos princípios ativos, formando um ciclo vicioso de alto custo econômico e ambiental.

A comunidade científica internacional, em grande parte responsável pelo desenvolvimento tecnológico gerador dos atuais problemas ambientais, foi obrigada a rever os seus conceitos, o que levou a uma mudança dos paradigmas norteadores da pesquisa agropecuária. Considera-se como agricultura sustentável aquela baseada na utilização racional dos recursos naturais visando o atendimento das necessidades das gerações presentes e futuras [72]. Em termos práticos, pretende-se que a dependência de insumos químicos por parte dos sistemas de produção agrícola seja diminuída, aumentando a importância dos processos biológicos na ciclagem de nutrientes e no controle biológico de pragas e doenças de plantas e animais domésticos. Estes processos são em sua maioria mediados e dependentes da ação microbiana em solos, águas e até em tecidos das plantas e animais. Esta é a razão da grande

importância do estudo e preservação da diversidade microbiana, tanto nos ecossistemas "naturais" (na ausência do distúrbio pelo homem) como nos agroecossistemas.

O descobrimento de *G. diazotrophicus* tornou-a um modelo pioneiro de bactéria endofítica fixadora de nitrogênio, abrindo caminho para estudos sobre fixação de nitrogênio em plantas não-leguminosas. Seu habitat endofítico, a capacidade de fixar nitrogênio na presença de nitratos [1], assim como sua capacidade de crescer *in vitro* em condições similares às encontradas em cana-de-açúcar [73] elevaram as expectativas sobre seu potencial agrobiotecnológico na realização de uma agricultura sustentável. Isto significa que ao multiplicar-se no interior do hospedeiro, esta bactéria enfrenta menor competição por nutrientes em comparação com o ambiente rizosférico e poderia dispôr de fontes abundantes de carbono, produtos da fotossíntese, para fixar o nitrogênio. O nitrogênio, produto da fixação biológica, seria excretado diretamente no interior dos tecidos da planta, que o utilizaria para suprir suas necessidades de síntese de macromoléculas essenciais. Este potencial biotecnológico foi ampliado pela demonstração que *G. diazotrophicus* efetivamente excreta 50% do nitrogênio fixado [74] e pela capacidade de produzir auxinas e citocinas [3, 67], as quais poderiam exercer efeito direto sobre a planta, produzir bacteriocina [65] e solubilizar compostos de zinco [68], como demonstrado anteriormente.

Decorre daí, portanto, o interesse em se estudar mais profundamente esta bactéria para que se possa conhecer seu potencial metabólico e suas características fisiológicas, ou seja, as relações genótipo metabólico / fenótipo. Assim sendo, incluímos neste trabalho estudos proteômicos e de bioinformática que vieram a complementar o projeto de seqüenciamento e anotação do genoma da *G. diazotrophicus* (projeto genoma Riogene) e revelar funções e mecanismos de ação bacteriana até então desconhecidos para a *G. diazotrophicus*.

1.2- Projetos Genoma

1.2.1- Estado da arte

Com a automatização dos métodos de seqüenciamento, e mais recentemente o desenvolvimento de métodos que permitem o seqüenciamento de um genoma bacteriano completo em apenas um dia [75, 76], seqüências completas dos genomas de centenas de microorganismos já estão disponíveis (<http://www.tigr.org>; <http://www.sanger.ac.uk>; <http://www.ncbi.nlm.nih.gov>). Em janeiro de 2007 estavam disponíveis os genomas completos de 312 bactérias e 26 archaea. Projetos genômicos mais complexos e completos, incluindo o humano, também já foram finalizados [77, 78].

Embora sejam automatizados, os métodos disponíveis não possibilitam o seqüenciamento do DNA inteiro de um organismo numa única reação. Dessa forma, faz-se necessária uma quebra randômica do DNA a ser seqüenciado, seguida do seqüenciamento de pequenos fragmentos gerados (“leituras”). Posteriormente esses fragmentos precisam ser montados utilizando-se algoritmos que percebam a sobreposição de seqüências de entre diversos fragmentos distintos, agrupando-os para formar fragmentos maiores (“contigs”) e assim sucessivamente até a obtenção de um fragmento único, o genoma. Depois de montada, a seqüência gerada precisa ser anotada. Anotar um genoma significa determinar por previsão computacional as possíveis regiões codificantes (ORFs) e atribuir possíveis funções a cada uma delas, através de análise por homologia de seqüências, buscas por domínios e estruturas protéicas características.

Dentre os membros da família *Acetobacteraceae*, *Granulibacter bethesdensis*, *Acidiphilium cryptum* e *Gluconobacter oxydans* 621H já tiveram seus genomas seqüenciados e anotados [79]. Os dois primeiros projetos foram completados no final do ano de 2006, de forma que embora os dados estejam disponíveis (<https://www.ncbi.nlm.nih.gov>), ainda não foram publicados.

G. oxydans é uma bactéria conhecida por oxidar incompletamente diversos carboidratos e alcoóis, num processo conhecido como fermentação oxidativa. Os produtos desses processos são, em grande parte, secretados para o meio, de forma que esta bactéria tem sido utilizada para produção de compostos com aplicação industrial [80, 81], tais como vitamina C e da droga anti-diabetes Miglitol [81]. O projeto genoma de *G. oxydans* revelou a presença de 2432 ORFs, numa sequência de 2.702.173 pares de base. Com esses dados, as vias metabólicas de oxidação incompleta que geram produtos importantes para a indústria foram elucidadas e diversas desidrogenases ligadas à membrana foram consideradas essenciais nestes processos [79].

Dentre os membros da família das Rhizobiaceae, que inclui as bactérias fixadoras de nitrogênio mais bem estudadas, algumas dezenas de espécies tiveram seus genomas seqüenciados, incluindo: *Rhizobium leguminosarum*, *Rhodopseudomonas palustris*, *Agrobacterium tumefaciens*, *Mesorhizobium sp*, *Rhizobium etli* e *Sinorhizobium meliloti* [82-84]. A finalização de tais genomas é essencial para estudos comparativos entre os diversos modos de interação entre bactérias diazotróficas e plantas.

1.2.2- O projeto Riogene

O Riogene é um instituto virtual formado por grupos de diversas universidades e instituições de pesquisa do Estado do Rio de Janeiro, com o intuito de estabelecer uma rede com competência para analisar genomas, formar cientistas nas áreas de genômica, biologia estrutural e bioinformática, assim como de seqüenciar genomas. Um desses grupos é a Unidade Multidisciplinar de Genômica, onde grande parte do trabalho descrito nesta dissertação foi realizado.

O primeiro projeto do Riogene, iniciado no ano de 2002, tem como objetivo seqüenciar e anotar o genoma completo da *G. diazotrophicus* (<http://www.lncc.br/~Riogene>). Em 2005, o Riogene anunciou o seqüenciamento de 99% do genoma de *G. diazotrophicus*.

No entanto, a sequência ainda encontrava-se dispersa em 455 *contigs*. Ainda assim, o processo de anotação do genoma foi iniciado pelos diversos grupos, utilizando o sistema de anotação de genomas bacterianos BioNotes, desenvolvido pelo próprio Riogene (tese de doutorado Melissa Lemos, PUC/RJ) [85]. Simultaneamente, estratégias de re-sequenciamento e PCR (reação em cadeia da polimerase), assim como novas ferramentas para montagem das sequências (dissertação de mestrado Marcelo Bertalan, IBqM/UFRJ) [86] foram utilizadas para finalizar o sequenciamento do genoma. A cada nova versão do genoma, as anotações das ORFs foram manualmente aperfeiçoadas. Com esses esforços, em agosto de 2006, o número de *contigs* foi reduzido a 8 somando um total de 2.995 ORFs, numa sequência de 3.954.735 pares de base. Em dezembro de 2006, o Riogene, juntamente com a FAPERJ, financiadora da rede, anunciaram a finalização do projeto genoma.

1.3- Projetos proteoma

1.3.1- Estado da arte

O sequenciamento dos genes que constituem o genótipo do organismo permite a dedução das seqüências de aminoácidos das proteínas, mas nada revela a respeito da expressão dos genes, que é um processo altamente complexo, incluindo o fato que a grande maioria das proteínas sofrem modificações pós-tradução. Informações sobre estas modificações, sobre a regulação da expressão, o funcionamento individual e/ou em conjunto das proteínas no organismo não podem ser deduzidas apenas das seqüências gênicas. Torna-se, portanto, necessário estudar o conjunto de proteínas expressas, que produzem o fenótipo do organismo. Surge assim a abordagem da Proteômica, que objetiva estudar a totalidade, ou conjuntos de proteínas no interior de uma célula (por exemplo: célula em funcionamento normal ou anormal ou submetida a tratamentos experimentais), permitindo identificar alterações e interações que ocorrem dentro de uma célula, no decorrer de um processo biológico.

Portanto, ao contrário do genoma, que é praticamente constante, o conjunto de proteínas expresso por um determinado organismo, ou seu proteoma, pode variar em função de condições ambientais [87]. Assim, o mesmo organismo pode apresentar diferentes proteomas, os quais são resultado da expressão de um conjunto de genes e das modificações pós-tradução das proteínas sintetizadas, em resposta às condições ambientais definidas [88]. A análise de proteomas tem como objetivo identificar as proteínas expressas por uma célula em um dado momento. A identificação de proteínas na escala proteômica, envolvendo dezenas, centenas ou mesmo milhares de proteínas, em função do estado fisiológico das células analisadas, poderá responder a questões fundamentais sobre os mecanismos biológicos muito mais rapidamente do que pela estratégia de análise de proteínas individuais [89]. O quadro global gerado pela análise de proteomas pode, por exemplo, permitir a construção de um mapa complexo de funções celulares, ao demonstrar como alterações em uma via de sinalização afetam outras vias ou quantas proteínas de uma via de sinalização interagem com outra(s) via(s). Este quadro global permite ainda conhecer a multiplicidade de fatores envolvidos em várias doenças, muito poucas das quais são causadas por um único gene, ou ainda identificar alvos de drogas, isto é, proteínas que seletivamente interagem com uma droga para produzir uma resposta definida [90]. A comparação da expressão gênica de cepas de microorganismos patogênicos com a das não patogênicas, por outro lado, pode ser usada com a finalidade de desenvolver métodos de diagnóstico ou terapêuticos [87, 91]. Prover detalhes quantitativos, detectar modificações pós-traducionais, identificar interações protéicas, estabelecer a dinâmica celular de proteínas e seus complexos, caracterizar respostas celulares a ambientes alterados, incluindo a presença de ligantes e comparar estados de doença e saúde são algumas das atribuições e desafios dos especialistas em proteômica hoje [92-94].

O método mais utilizado para a caracterização inicial de um proteoma é a eletroforese bidimensional (2D-GE). A habilidade de resolver e investigar a abundância de muitas centenas de proteínas de uma mesma amostra é sua principal característica. Essa técnica combina focalização isoeletrica (primeira dimensão) numa tira de gradiente de pH imobilizado, seguida pela eletroforese em gel de poliacrilamida (segunda dimensão), separando centenas de polipeptídeos pelos seus pontos isoeletricos (pI) e massas moleculares (MM) [88, 95-97]. Assim, o pI e a MM de cada banda no gel 2D podem ser determinados experimentalmente e a identificação de cada proteína pode ser obtida posteriormente por espectrometria de massas, seja análise de mapa peptídico após hidrólise ou micro-seqüenciamento e buscas por homologia em bancos de dados de seqüências [98, 99].

A introdução dos gradientes imobilizados de pH, que permitem uma maior reprodutibilidade dos géis, ampliou o uso da 2D-GE para análise de sistemas complexos. Além disso, as técnicas de micro-seqüenciamento de peptídeos possibilitam a obtenção de uma seqüência de cerca de dez aminoácidos a partir do material presente em uma banda de um gel 2D [100, 101].

As proteínas de interesse, resolvidas no gel bidimensional, podem ser recortadas do gel e digeridas usando uma protease altamente específica (geralmente a tripsina); e os peptídeos derivados podem ser eluídos do gel e analisados por espectrometria de massas. O problema da identificação da proteína foi então reduzido à identificação dos peptídeos. Entre outras, as vantagens dessa metodologia são: 1) a facilidade com que os peptídeos são obtidos pela digestão e eluídos do gel, comparados às proteínas e 2) o maior poder de resolução de moléculas menores na espectrometria de massa.

No espectrômetro, os peptídeos são ionizados, acelerados e detectados (fig. 1 do anexo) [102]. Na técnica MALDI-TOF, a fonte de ionização é um laser (geralmente o comprimento de onda é de 337 nm). Neste caso, a amostra é pré-misturada a uma matriz

orgânica, capaz de absorver a energia no comprimento de onda emitido pelo laser utilizado. Quando o laser é incidido na amostra, a energia absorvida é utilizada tanto para ionizar os peptídeos, através da transferência de prótons da matriz para os peptídeos, como para desorver estas moléculas ionizadas em uma câmara submetida a uma grande diferença de potencial elétrico. Os íons são acelerados até entrarem numa região livre de campo, o tubo de vôo. O analisador do tipo tempo de vôo (TOF) permite obter a relação massa/carga do peptídeo. Tendo a mesma carga e a mesma quantidade de movimento ao final da aceleração, os íons mais pesados demoram mais tempo para percorrer o tubo do que os mais leves. O tempo de vôo para cada íon detectado é plotada num espectro podendo-se deduzir para cada pico uma relação massa/carga. Quando dois analisadores de massa, por exemplo dois TOFs, são colocados em seqüência no aparelho (MALDI-TOF/TOF), torna-se possível, em uma primeira etapa, selecionar um único peptídeo da amostra e induzir, em uma câmara contendo um gás inerte, sua fragmentação, que ocorre preferencialmente nas ligações peptídicas, gerando fragmentos cuja diferença de massa, detectada em uma segunda etapa, possibilita a dedução da seqüência de aminoácidos do peptídeo analisado. Através de buscas em bancos de dados a proteína pode ser identificada [103, 104].

Recentemente, a cromatografia líquida (LC) de alta performance tem sido amplamente empregada para separação de amostras complexas [105-109]. Em substituição à eletroforese bidimensional, a separação de amostras em colunas cromatográficas, permite a separação inclusive de proteínas hidrofóbicas, proteínas de alto e baixo peso molecular, assim como proteínas com valores de pI extremos, as quais não são resolvidas por 2D-GE. Em alguns casos, uma mistura complexa de peptídeos é digerida por uma enzima específica e os peptídeos resultantes são separados em uma ou mais colunas. Neste caso, a identificação da proteína requer o seqüenciamento de alguns de seus peptídeos através da fragmentação e o uso seqüencial de dois analisadores de massas (LC-MS/MS).

Geralmente, esta estratégia, conhecida como MudPit (*Multi-Dimensional Protein Identification Technology*) [110], inclui a utilização de espectrômetros de massa com uma fonte de ionização do tipo *eletronspray* (ESI) e dois analisadores de massa, como por exemplo um triplo quadrupolo e um por tempo de voo.

As amostras a serem ionizadas por ESI devem ser solúveis em solventes de baixo ponto de ebulição, tais como água, metanol e acetonitrila, e estáveis a baixas concentrações. O sistema consiste em um capilar ou agulha extremamente fina (geralmente 75 a 150 μm) e uma série de colimadores (fig. 2 do anexo). A solução a ser analisada flui pela agulha. Na ponta do capilar, situado na fonte de ionização do espectrômetro de massa, é aplicada uma voltagem de 3 a 4 kV. Em consequência do forte campo elétrico gerado, a amostra que emerge na ponta do capilar é dispersada em um aerosol de gotas altamente carregadas. Um gás nebulisante, geralmente o nitrogênio, flui co-axialmente à parede externa do capilar, direcionando o fluxo do *spray* para o interior do espectrômetro de massas. As gotas carregadas diminuem de tamanho pela evaporação do solvente, de forma que forças coulombianas induzem instabilidade da superfície da gota, gerando uma pequena explosão que resulta na liberação de íons multicarregados do analito [104].

Em um espectrômetro de massas, os analisadores compõem a parte do equipamento que permite a mensuração da relação massa/carga (m/q) dos íons. Assim, para obtenção da relação m/q , existem instrumentos com base nas trajetórias dos íons em um campo magnético (instrumentos de "setor magnético"); os que usam o tempo necessário para que íons com determinada energia cinética percorram um determinado percurso (instrumentos de "tempo de voo"), ou outros que aproveitam a estabilidade dependente da relação m/q em campos quadrupolares oscilantes (instrumentos do tipo "quadrupolo"). Além disso, foram desenvolvidos equipamentos capazes de confinar os íons em uma determinada região do espaço e medir a relação m/q : os espectrômetros de massas denominados "armadilha de íons

quadrupolar", e os espectrômetros de massas que analisam as órbitas (frequências) dos íons confinados por "Transformada de Fourier". Em cada uma destas versões existem variações em termos da formação dos íons, montagem do instrumento e sistemas de detecção de acordo com o interesse de aplicação.

Entre os primeiros proteomas bacterianos publicados encontram-se os estudos com *Mycoplasma genitalium* e abordagens comparativas dos patógenos *Mycobacterium tuberculosis* e *Mycobacterium bovis*, linhagens BCG [111, 112]. Desde então, diversos processos metabólicos em microorganismos já foram analisados por estas estratégias [113-116], incluindo algumas iniciativas envolvendo bactérias simbiotes e fixadoras de nitrogênio [117-120].

Em 1999, Guerreiro e colaboradores [119] estabeleceram um mapa de eletroforese bidimensional de proteínas de *Sinorhizobium meliloti*, contendo aproximadamente 2000 produtos gênicos, na faixa de pH entre 4 e 7 e de peso molecular entre 10 e 122 kDa. Considerando que o genoma de *S. meliloti* possui 6,5 Mpb, potencialmente codificando 6.500 proteínas, 31% da capacidade codificante de *S. meliloti* foi visualizada em gel 2D. Dezoito proteínas constitutivas abundantes tiveram seu N-terminal seqüenciado e diversas proteínas ainda não conhecidas em rizóbios foram identificadas, incluindo componentes de um sistema de transporte ativo para captação de amidas e uréia. Ainda neste trabalho, foram identificadas proteínas induzidas ou reprimidas durante a transição de fases da curva de crescimento, tais como uma pequena proteína de choque térmico, envolvida na proteção contra degradação, e uma hidroxilase tipo glioxilase II, envolvida na detoxificação de metabólitos tóxicos. Cinco proteínas envolvidas da tradução e enovelamento tridimensional foram encontradas com mobilidade eletroforética e isoformas idênticas às suas homólogas em *Rhizobium loti*, indicando que essas proteínas são altamente conservadas no gênero *rhizobium*.

Em 2003, Djordjevic e colaboradores [120] continuaram a caracterização das proteínas de *S. meliloti* e identificaram 1180 proteínas, produtos de 810 genes, utilizando amostras da bactéria crescida em diferentes meios de cultivo e recuperadas de nódulos simbióticos. Neste estudo, a atividade de 53 vias metabólicas foi inferida, 146 proteínas de membrana e diversos transportadores e alguns reguladores foram detectados. 27 proteínas foram expressas especificamente em nódulos e outras 35 em resposta ao estresse de nutrientes. Ainda, foram identificadas 7 novas proteínas, não previstas a partir da sequência de DNA.

A análise de proteínas bacterianas por eletroforese bidimensional seguida de espectrometria de massas também foi realizada em *Bradyrhizobium japonicum* [121], *Agrobacterium tumefaciens* [122], *Acetobacter aceti* [123], *Brucella melitensis* [124], *Bacillus sp* [125], *Escherichia coli* [126], *Staphylococcus aureus* [127], *Pseudomonas aeruginosa* [128], *Streptococcus mutans* [129], *Vibrio cholerae* [113] entre outros. Todos esses trabalhos têm contribuído bastante para o melhor entendimento da regulação de vias metabólicas, direcionando estudos com abordagens mais racionais e específicas.

1.3.2- Rede Proteômica do Rio de Janeiro (PROTEOMA-RIO)

A Rede Proteômica do Rio de Janeiro (PROTEOMA-RIO) foi criada por iniciativa da FAPERJ, sendo constituída por um conjunto de laboratórios localizados em diferentes instituições do Rio de Janeiro, equipados com tecnologias complementares e dedicados à caracterização de proteomas e identificação de proteínas individuais. Um desses laboratórios é a Unidade Multidisciplinar de Genômica do IBCCF, na qual grande parte desta dissertação foi realizada.

Inicialmente quatro projetos foram selecionados para execução pela rede PROTEOMA-RIO: análise proteômica da bactéria endofítica *G. diazotrophicus* e de sua interação com plantas de cana-de-açúcar, regulação da expressão gênica em *Vibrio cholerae*, estudo proteômico da infecção pelo vírus da dengue e de toxinas animais e seus inibidores

naturais. Posteriormente outros projetos foram adotados pelo PROTEOMA-RIO, entre eles o estudo de fungos patogênicos e análise proteômica diferencial de leucemias.

Esta iniciativa já resultou na publicação de pelo menos 3 trabalhos envolvendo o grupo da Unidade Multidisciplinar de Genômica: um mapa de referência de proteínas de *Vibrio cholerae* [91], análise proteômica diferencial de *Vibrio cholerae* em abundância e limitação de fosfato no meio de cultura [113] e identificação de marcadores tumorais para leucemia mielóide crônica [130].

O projeto da *G. diazotrophicus* iniciou-se com a análise comparativa da cepa PAL5 crescida em limitação e abundância de compostos nitrogenados, em fase exponencial ou estacionária de crescimento. Uma análise preliminar destes dados foi apresentada na minha monografia de final de curso de graduação [131]. Um outro grupo da rede PROTEOMA-RIO está investigando proteínas de *G. diazotrophicus* expressas especificamente em resposta à presença da planta de cana-de-açúcar e proteínas envolvidas na formação de biofilme.

A continuação do projeto proteoma da *G. diazotrophicus* será apresentada ao longo desta dissertação de mestrado.

2- OBJETIVOS



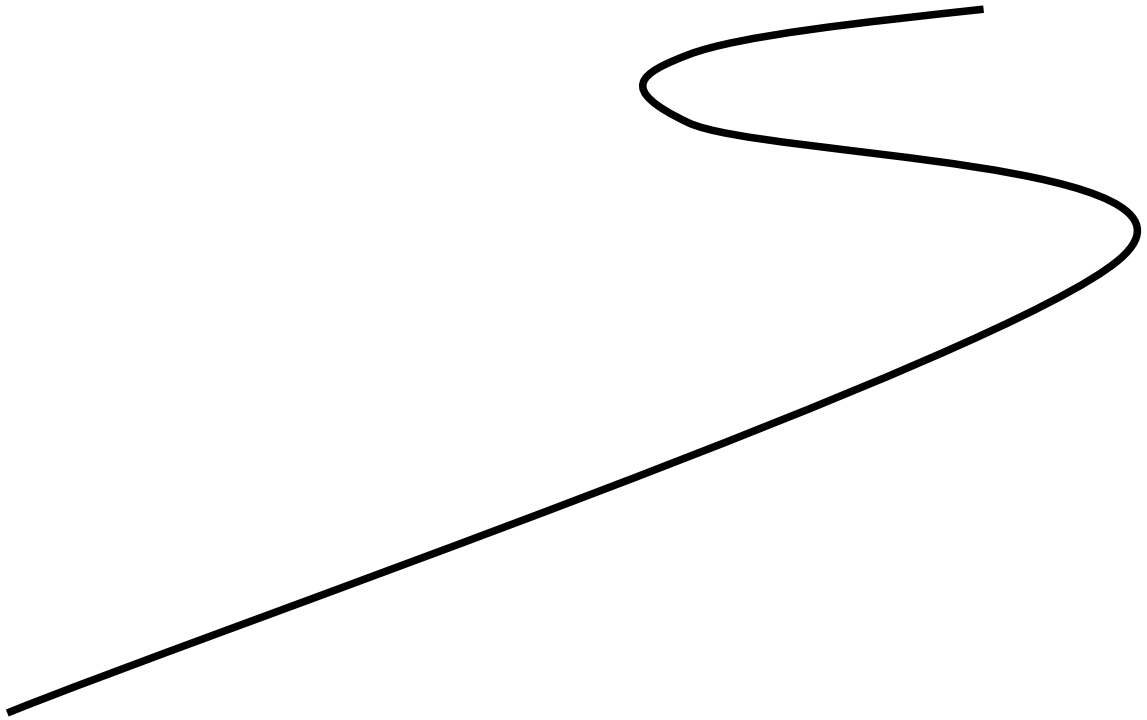
2- OBJETIVOS

Identificar o maior número possível de proteínas de *G. diazotrophicus*, de forma a contribuir para o melhor entendimento da fisiologia desta bactéria.

2.1- Objetivos específicos

- Estabelecer um banco de proteínas de *G. diazotrophicus*: um conjunto de proteínas expressas, suas características (tais como massa molecular e ponto isoelétrico) e vias metabólicas a que pertencem;
- Análise proteômica diferencial de *G. diazotrophicus* crescida em limitação e abundância de compostos nitrogenados;
- Relacionar os dados obtidos por proteômica com os dados do projeto genoma da *G. diazotrophicus* (Riogene – em andamento), validando os dados do seqüenciamento e atribuindo informações sobre as proteínas identificadas.

3- MATERIAIS E MÉTODOS



3.1- Cepa e condições de cultivo

3.2- Lise celular

3.3- Separação de proteínas

3.4- Análise de imagens

3.5- Espectrometria de massas

3.6- Identificação de proteínas

3.7- Análise de bioinformática

3- MATERIAIS E MÉTODOS

3.1- Cepa e condições de cultivo

G. diazotrophicus PAL5 foi cultivada no Laboratório de Genética e Bioquímica da Embrapa Agrobiologia, sob orientação da pesquisadora Kátia Teixeira. Culturas da cepa *G. diazotrophicus* PAL5 (BR11281, banco de microorganismos EMBRAPA Agrobiologia) foram obtidas em 100 ml de meio LGI (K_2HPO_4 0,2g/L, KH_2PO_4 0,6g/L, $MgSO_4 \cdot 7H_2O$ 0,2g/L, $CaCl_2 \cdot 2H_2O$ 0,02g/L, $Na_2MoO_4 \cdot 2 H_2O$ 0,002g/L, $FeCl_3 \cdot 6 H_2O$ 0,01g/L, 5ml/L de azul de bromotimol 0,5% em 0,2M de KOH, glicose 10g/L) contendo 10 mM de sulfato de amônia, $(NH_4)_2SO_4$. Na fase exponencial do crescimento, alíquotas desta cultura foram inoculadas em meios de cultura frescos, na proporção 1:1000, para obtenção de massa celular para as análises proteômicas.

Para compor o banco proteômico de *G. diazotrophicus*, estas culturas foram submetidas a duas condições diferentes: a) em um fermentador em meio LGI e, b) em recipientes *erlenmeyers* em meio DYGS. Na condição “a”, as bactéria eram cultivadas em LGI suplementado com 10mM de $(NH_4)_2SO_4$ e pH constante em 5,5, simulando a situação vivenciada pela bactéria na planta, conforme descrito por Cavalcante e Dobereiner (1988) [1] e Stephan e colaboradores (1991) [21]. Para estes crescimentos foi utilizado o Fermentador Bioengineering, AG – capacidade para 2,5 L, como recomendado por Dobereiner e colaboradores, 1995 [18]. Nestas condições, o fluxo de oxigênio, a temperatura de 30°C e pH de 5,5 (tampão citrato-fosfato, injetado na cultura por bombas peristálticas, acionadas pela alteração do pH do meio) foram mantidos constantes. O controle da contaminação foi feito com filtros, conectados às vias de entrada do fluxo de gases e a outras aberturas utilizadas para o controle da pressão interna do vaso do fermentador. As amostras foram coletadas em frascos estéreis e o crescimento foi acompanhado pela contagem do número de células em câmara de Neubauer e determinação da absorbância em comprimento de onda de 600 nm no

espectrofotômetro (PE - Modelo Lambda 11). A presença ou consumo de N, assim como os valores de pH e O₂ dissolvido, também foram acompanhados ao longo do crescimento. A temperatura de incubação foi mantida em 30°C, com o auxílio de banho-refrigerado recirculante, e a agitação foi fixada em 400 rpm. A coleta das amostras foi feita no início da fase estacionária do crescimento de *G. diazotrophicus*.

Na condição “b”, também para obtenção de células para o banco proteômico, as células foram crescidas em meio DYGS (glicose 2 g/l, extrato de levedura 2 g/l, peptona 1,5 g/l, ácido glutâmico 1,3 g/l, K₂HPO₄ 0,5 g/l, MgSO₄.7H₂O 0,5 g/l, pH 6,0), em *erlenmeyers*, à temperatura e agitação contantes de 30°C e 180 rpm. Neste caso, a cultura foi centrifugada e as células coletadas na fase exponencial do crescimento de *G. diazotrophicus* para o preparo do lisado.

Para a análise proteômica diferencial, alíquotas das pré-culturas em LGI foram lavadas 2 vezes e ressuspensas em meio LGI sem sais nitrogenados antes de serem inoculadas no meio de cultivo. As células foram crescidas em fermentador, em LGI suplementado com 1 ou 10 mM de (NH₄)₂SO₄. O acompanhamento do crescimento, assim como o controle do fluxo de oxigênio foram monitorados conforme descrito acima. A temperatura foi mantida a 30°C, a agitação a 400 rpm e o pH inicial foi de 5,5, mas variou durante o crescimento.

3.2- Lise celular

Extratos citoplasmáticos celulares para análise por eletroforese bidimensional foram obtidos pela adição de um tampão de lise (Uréia 8M, Tris 40 mM, fluoreto de metil fenil sulfonato (PMSF) 1 mM, “Pharmalyte” 0,5% e CHAPS 4% ou Triton 2%) ao sedimento de células (na proporção de 500 µl de tampão para o sedimento de uma suspensão de células, de densidade ótica em 600nm igual a 2,0). As suspensões foram submetidas a cerca de 20 ciclos de congelamento em N₂ líquido e aquecimento a 37°C e a lise foi acompanhada em microscópio ótico. Quando necessário, a suspensão foi submetida ciclos adicionais até a lise

estar completa. Os lisados foram centrifugados a 14.000 x g por 60 minutos e o sobrenadante contendo as proteínas solúveis foi coletado e estocado a -80°C ou em N₂ líquido. Alíquotas de cada extrato foram submetidas a eletroforese em gel unidimensional após serem quantificadas pelo método de Bradford [132] ou com o 2D Quant Kit (GE Healthcare).

Para análise pela técnica de LC-MS/MS, os extratos celulares foram obtidos pela adição do tampão de lise descrito por Laemmli (Tris-Cl pH 6,8 100 mM, ditioneitol 200 mM, dodecilsulfato de sódio 4%, azul de bromofenol 0,2%, glicerol 20%) [133] ao sedimento de células. Esses extratos foram fervidos por 5 minutos e centrifugados por 30 minutos a 14000 rpm.

3.3- Separação de proteínas

As proteínas dos lisados celulares obtidos para o banco proteômico foram analisadas por eletroforese bidimensional, no Sistema IPGPhor (GE Healthcare). Para a primeira dimensão, tiras de 18 cm e pH 4-7 imobilizado em poliacrilamida (Immobiline DryStrip) foram utilizadas e a focalização foi realizada de acordo com as instruções do fabricante. Para a segunda dimensão, utilizou-se um gel SDS-PAGE 12% e o sistema DALTsix, ambos da GE Healthcare. Antes da focalização isoeletrica 1 mg da amostra foi solubilizada em 240 µl de solução de reidratação (Uréia 8M, Pharmalyte 0,5%, Chaps ou Triton X-100 2%, traços de azul de bromofenol). Após a focalização, as tiras foram equilibradas em 10 ml de soluções de equilíbrio (Uréia 6M, Tris-HCl 50mM pH 8,8, Glicerol 30%, dodecilsulfato de sódio (SDS) 2%, azul de bromofenol), contendo ditioneitol (DTT) 0,01%, como agente redutor, e iodoacetoamida 0,025% como agente alquilante, por 15 minutos cada, a temperatura ambiente e agitação suave.

No caso dos extratos celulares obtidos para análise proteômica diferencial, a focalização isoeletrica foi realizada em tiras de 18 cm e pH imobilizado 3-10; e a segunda dimensão foi feita em gel SDS-PAGE 12-14%, adquiridos da GE Healthcare. Ambas

dimensões foram corridas no aparelho MultiphorII (GE Healthcare) e seguiram instruções do fabricante. Neste caso, 500µg de amostra foram dissolvidas em solução de reidratação. As etapas de equilíbrio após a focalização ocorreram conforme descrito acima e o padrão de baixo peso molecular (GE Healthcare) foi utilizado na segunda dimensão para determinação relativa das MM de cada *spot*.

Após a segunda dimensão, os géis foram submetidos a tratamentos de fixação em 30% de etanol e 0,5% ácido acético por 12 a 16 horas, coloração por *Coomassie Blue* R250 0,02% em 30% de etanol por 3 horas e descoloração com 30% de etanol. Os géis corados foram escaneados no ImageMaster LabScan v3.01 (com resolução de 300 dpi) e analisados através do programa ImageMaster Platinum v1.0 (GE Healthcare).

As proteínas dos extratos celulares para análise por LC-MS/MS (banco proteômico) foram previamente separadas por eletroforese unidimensional SDS-PAGE, em gel de acrilamida 12%, no sistema mini-Protean II (Bio-Rad). O gel foi corado por *Coomassie Blue* R250 e a pista contendo a amostra foi dividida em 10 partes iguais ao longo do comprimento.

3.4- Análise de imagens

Toda a análise das imagens foi feita pelo programa ImageMaster Platinum, conforme recomendações do manual [134]. O primeiro passo na análise dos géis foi a detecção automática dos *spots* no programa ImageMaster Platinum seguida de uma análise manual mais apurada, identificando *spots* não detectados pelo programa. *Spots* duplos considerados como único e falsos *spots* foram identificados e corrigidos. A seguir foram utilizadas as ferramentas de remoção de *background* e normalização. Posteriormente, os géis foram calibrados, utilizando uma escala linear de pH e padrões de peso molecular (GE Healthcare).

Para a análise das proteínas diferencialmente expressas foram obtidas amostras de quatro condições: 1) fase exponencial (E) / limitação de compostos nitrogenados (L), EL; 2) fase exponencial (E) / abundância de compostos nitrogenados (H), EH; 3) fase estacionária

(S) / limitação de compostos nitrogenados (L), SL; 4) fase estacionária (S) / abundância de compostos nitrogenados (H), SH. Para cada condição foram obtidos géis em triplicatas, sendo duas amostras oriundas de uma mesma cultura de células e a terceira de um cultivo diferente. Cada gel do conjunto de triplicatas foi denominado de “gel membro”.

Os 3 géis membros foram sobrepostos e um “gel master” foi obtido para cada condição, contendo apenas os *spots* comuns aos géis membros, ou seja, apenas *spots* reprodutíveis. A posição e densidade ótica dos *spots* nos géis “master” é uma média desses valores para o *spot* nos géis membros. *Spots* não presentes nos 3 géis membro ou cujo volume relativo diferiu mais de 20% entre eles, não foram considerados para as análises posteriores.

Com o intuito de identificar proteínas específicas das condições em estudo (limitação x abundância de compostos nitrogenados; fase exponencial x estacionária de crescimento), interseções dos géis “master” foram obtidas e comparadas. *Spots* comuns a dois géis “master” foram considerados diferencialmente expressos quando seu volume relativo variou em mais de 3 vezes. O volume relativo (% Vol) é um valor normalizado do volume, que o torna relativamente independente de variações entre géis, geralmente causadas pelas condições experimentais. O cálculo do volume relativo está expresso na equação 1.

$$\%Vol = \frac{Vol}{\sum_{S=1}^n Vol_s} \times 100$$

Equação 1: Volume relativo (% Vol) de um *spot* expresso como o volume do *spot* multiplicado por 100 e dividido pela soma dos volumes de todos os *spots* do gel.

3.5- Espectrometria de massas

As proteínas visualizadas no gel do banco proteômico e as selecionadas pela análise diferencial foram digeridas em gel, pela enzima tripsina (Trypsin sequencing grade; Promega Madison, WI, USA), conforme o protocolo descrito pela UCSF Mass Spectrometry Facility (www.ucsf.edu) com as alterações observadas por von Krüger e colaboradores [113]. Os

peptídeos foram misturados com uma matriz de ácido α -cyano 4-hydroxicinâmico (CHCA) em solução 50% de acetonitrila (ACN) e 1% de ácido trifluoroacético (TFA) e analisados no aparelho Voyager DE PRO Biospectrometry Workstation (Applied Biosystem). Os espectros foram adquiridos no modo refletido com o dispositivo “extração atrasada”, na faixa massa/carga de 800 a 4000 daltons; analisados no programa Data Explorer v5.1, através do qual foram realizadas calibrações interna (utilizando-se como referências os picos 2211.1046, 1045.5642 e 842.5 da tripsina) e externa (utilizando como padrão uma mistura comercial de peptídeos de massa conhecida: Calmix2, Applied Biosystem), correção da linha de base, filtragem do ruído, seleção dos picos monoisotópicos e listagem dos picos bem resolvidos. A acurácia das massas obtidas nos espectros considerados aceitáveis foi melhor que 100 ppm. As buscas nos bancos de dados foram feitas pela interface do programa MS-FIT do Protein Prospector [135].

Nos casos em que a identificação da proteína não foi possível através da análise do espectro de digestão triptica ou mesmo para confirmar identificações por *fingerprint* de massa, as amostras foram posteriormente analisadas por espectrometria de massa MALDI-TOF/TOF (4700 Explorer Proteomics Analyzer, Applied Biosystem), do Laboratório de proteoma do departamento de fisiologia e farmacodinâmica da Fundação Oswaldo Cruz do Rio de Janeiro ou no Laboratório nacional de luz síncrotron (LNLS). Nesta técnica, para a dissociação por colisão do íon precursor, foi utilizado o N_2 como gás indutor, a pressão de 2.8×10^{-6} torr. As massas dos picos do espectro resultante, representantes dos íons “fragmento” gerados, devidamente calibradas, foram utilizadas para buscas no banco de dados através do programa Mascot [136] (Matrix Science), localmente instalado.

Por outro lado, nos casos que não obtivemos espectros de massa com intensidade suficiente para distinguir o sinal do ruído, utilizamos colunas de C18 para remover possíveis contaminantes, por exemplo sais, que afetam a análise dos peptídeos. Essas colunas foram

montadas no laboratório, utilizando *C18 Empore Extraction Disks* (3M[®]). Pequenos pedaços desses discos de C18 foram cortados com a ponta de um capilar e colocados dentro de uma ponteira de pipeta automática P20. Com o mesmo capilar, o material foi forçado até a ponta da ponteira, compactando-o (aproximadamente 1 a 2 mm de C18). Adaptamos a ponta de uma seringa ao encaixe da parte superior da ponteira, de forma que fosse possível ao apertar a seringa, fazer pressão sobre a coluna. A coluna foi equilibrada com 20 µl de uma solução 0,1% de TFA. Em seguida, a mistura de peptídeos trípticos eluídos do gel foi passada pela coluna e esta foi lavada com 20 µl de 0,1% de TFA. Os peptídeos foram eluídos da coluna na solução saturada da matriz CHCA em 50% ACN e 1% TFA. Este procedimento foi aprendido durante o curso de proteoma ministrado pelo Dr. Peter Roepstorf na Fundação Oswaldo Cruz do Rio de Janeiro, em fevereiro de 2006 [137].

Os 10 pedaços do gel contendo as proteínas separadas por eletroforese 1D a serem analisados por LC-MS/MS foram tratados com tripsina de acordo com o protocolo descrito acima. Os peptídeos trípticos dessas amostras foram analisadas no espectrômetro de massa Q-Tof Ultima API (Waters) do Laboratório Nacional de Luz Síncrotron (LNLS) cuja fonte de ionização é do tipo *eletronspray* e os analisadores são um quadrupolo e um tempo de vôleo. Este equipamento está acoplado a uma coluna de desalinização (Waters Opti-Pak C18 trap column) e uma coluna capilar C18 (Nanoease C18, Waters). Neste caso, a identificação de proteínas é realizada por seqüenciamento de peptídeos.

3.6- Identificação de proteínas

Tendo em vista que a identificação de proteínas por espectrometria de massa requer uma busca em bancos de dados, que o número de seqüências de aminoácidos da bactéria *G. diazotrophicus* em bancos de dados públicos é bastante restrito (um total de 66 seqüências depositadas no NCBI – 2006) e que o genoma desta bactéria está em fase de anotação (pelo Projeto Riogene), para a identificação de proteínas por mapa peptídico foi instalado

localmente o programa Mascot [136], e a ele acoplado versões do banco de dados das ORFs - calculadas pelo programa *Glimmer* [138] - dos *contigs* gerados pelo Projeto Riogene ou mesmo as sequências de DNA dos *contigs*. Proteínas do gel do banco proteômico foram identificadas no banco de dados do projeto Riogene. Proteínas da análise diferencial foram identificadas no banco do Riogene e ainda no banco de dados de α -proteobactérias do NCBI.

Em todas as buscas realizadas no programa Mascot foi utilizado o modo de busca por identidade e foi considerado o incremento de massa resultante da alquilação das cisteínas pela iodoacetoamida durante a etapa de equilíbrio das tiras isoeletrofocalizadas.

Como padrão para a identificação de uma proteína por *fingerprint* foram utilizados os critérios abaixo:

- Pontuação maior do que 55 ($p < 0,05$);
- Único *hit* com $p < 0,05$;
- Mínimo de 4 peptídeos do mapa peptídico presentes na seqüência do *hit*;
- Mínimo de 25% da seqüência da proteína *hit* coberta por peptídeos do mapa peptídico;

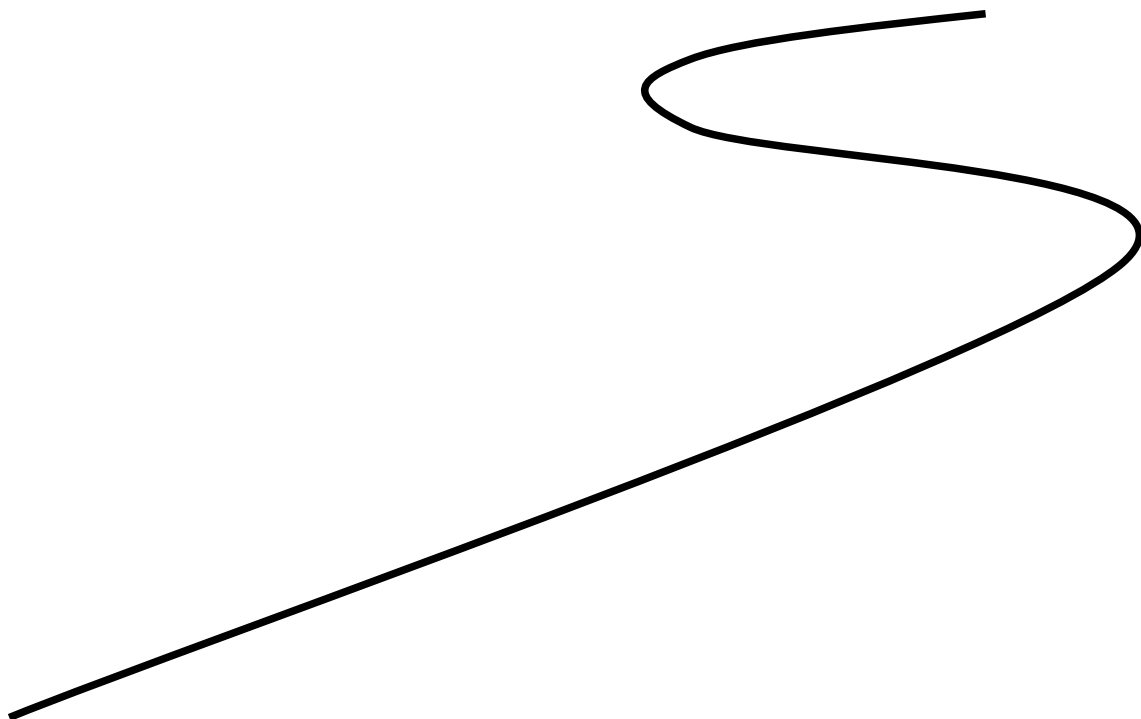
Nos casos em que foi possível obter sequências com mais de 5 aminoácidos para peptídeos, não foi necessário o preenchimento de todos os critérios acima descritos.

3.7- Análise bioinformática

Cada proteína identificada teve sua seqüência anotada, ou seja, foi realizada uma análise da estrutura e função do gene codificante de cada proteína. Tais análises foram realizadas utilizando as seguintes ferramentas de bioinformática: BLAST (busca por seqüências homólogas) [139, 140], PFAM (busca de domínios funcionais)[141], COG (busca de seqüências ortólogas) [142], KEGG (identificação de vias metabólicas) [143], TMHMM (previsão de hélices transmembranares) [144], PSORTb (previsão de localização celular) e Signalp (previsão de peptídeo sinal) [145]. Ainda, foi utilizado o programa Artemis (Sanger

institute) [146] e suas ferramentas para avaliação de início e final de ORFs, conteúdo de GC, avaliação de frames de leitura e da frequência de utilização de códons (*codon usage*) das ORFs. Em agosto de 2006, o genoma da *G. diazotrophicus*, de acordo com o projeto Riogene, apresentou aproximadamente 3000 ORFs, distribuídas em 8 *contigs*.

4 – RESULTADOS E DISCUSSÃO



4.1- Banco proteômico

4.2- Análise diferencial

4.3- Proteoma x Genoma

4.1- BANCO PROTEÔMICO

4.1- BANCO PROTEÔMICO

4.1.1 – Crescimento celular e obtenção das amostras

G. diazotrophicus foi cultivada em fermentador, no meio LGI, com pH constante em 5,5 por adição de tampão citrato, a temperatura de 30°C, no Laboratório de Genética e Bioquímica da EMBRAPA Agrobiologia. A amostra desta cultura utilizada para a obtenção de um mapa proteômico de *G. diazotrophicus* foi coletada em 120 horas (início da fase estacionária) (fig. 6).

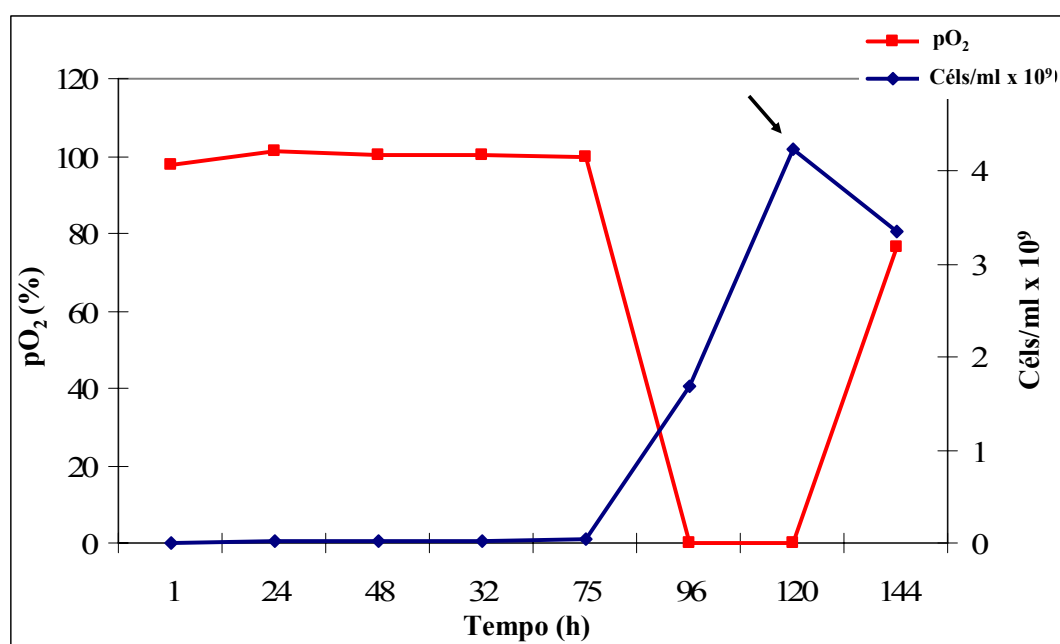


Figura 6: Curva de crescimento de células de *G. diazotrophicus* em meio LGI suplementado com 10mM de sulfato de amônio e controle do pH do meio de cultura. Em azul: número de células x 10⁹ por ml de cultura; em vermelho: pressão parcial de oxigênio (pO₂) dissolvido no meio de cultura; seta: ponto de coleta da amostra.

Além disto, *G. diazotrophicus* foi cultivada em meio DYGS, em *erlenmeyers*, sem controle do pH do meio de cultura, a temperatura de 30°C e agitação de 180 rpm. Neste caso, as amostras para o mapa proteômico foram coletadas após 76 horas de cultivo.

4.1.3- Construção do banco de proteínas de *G. diazotrophicus*

A construção do banco proteômico de *G. diazotrophicus* foi realizada em duas etapas. Na primeira, as proteínas foram separadas por 2D-GE e identificadas por espectrometria de massas MALDI-Tof ou Tof/Tof. Na segunda, utilizamos a estratégia de MudPit, na qual as proteínas foram digeridas por tripsina e os peptídeos resultantes separados em cromatografia e analisados por espectrometria de massas ESI / Q-Tof.

4.1.3.1- Eletroforese bidimensional e espectrometria de massas MALDI

O perfil protéico obtido por eletroforese bidimensional das amostras de *G. diazotrophicus* crescida em meio LGI suplementado com 10mM de $(\text{NH}_4)_2\text{SO}_4$ pode ser observado na figura 7. O banco proteômico contém 354 *spots* bem distribuídos ao longo de toda a faixa de pH 4 a 7. Cerca de 80% das proteínas visualizadas apresentaram peso molecular entre 22 e 97 kDa. Apenas 7% apresentaram MM maior do que 97 e 13% menor do que 22 kDa.

A concentração do gel utilizado não permite a visualização de proteínas de alto (> 120 kDa) e baixo (< 10 kDa) pesos moleculares, sendo portanto necessário a utilização de técnicas complementares para observação de tais proteínas. O Laboratório de Hemostase e Venenos, coordenado pela prof. Russolina Zingali, e mais recentemente a Unidade de Espectrometria de Massas e Proteômica (UEMP) possui ampla experiência na análise de peptídeos. Recentemente iniciamos uma colaboração com a prof. Russolina Zingali e Ana Lúcia de Oliveira-Carvalho (UEMP) para análise dos peptídeos de *G. diazotrophicus*. O objetivo deste projeto é analisar peptídeos (1 a 10 kDa) citoplasmáticos e peptídeos secretados para o meio de cultura. Contudo, os resultados ainda são preliminares e não serão apresentados nesta dissertação.

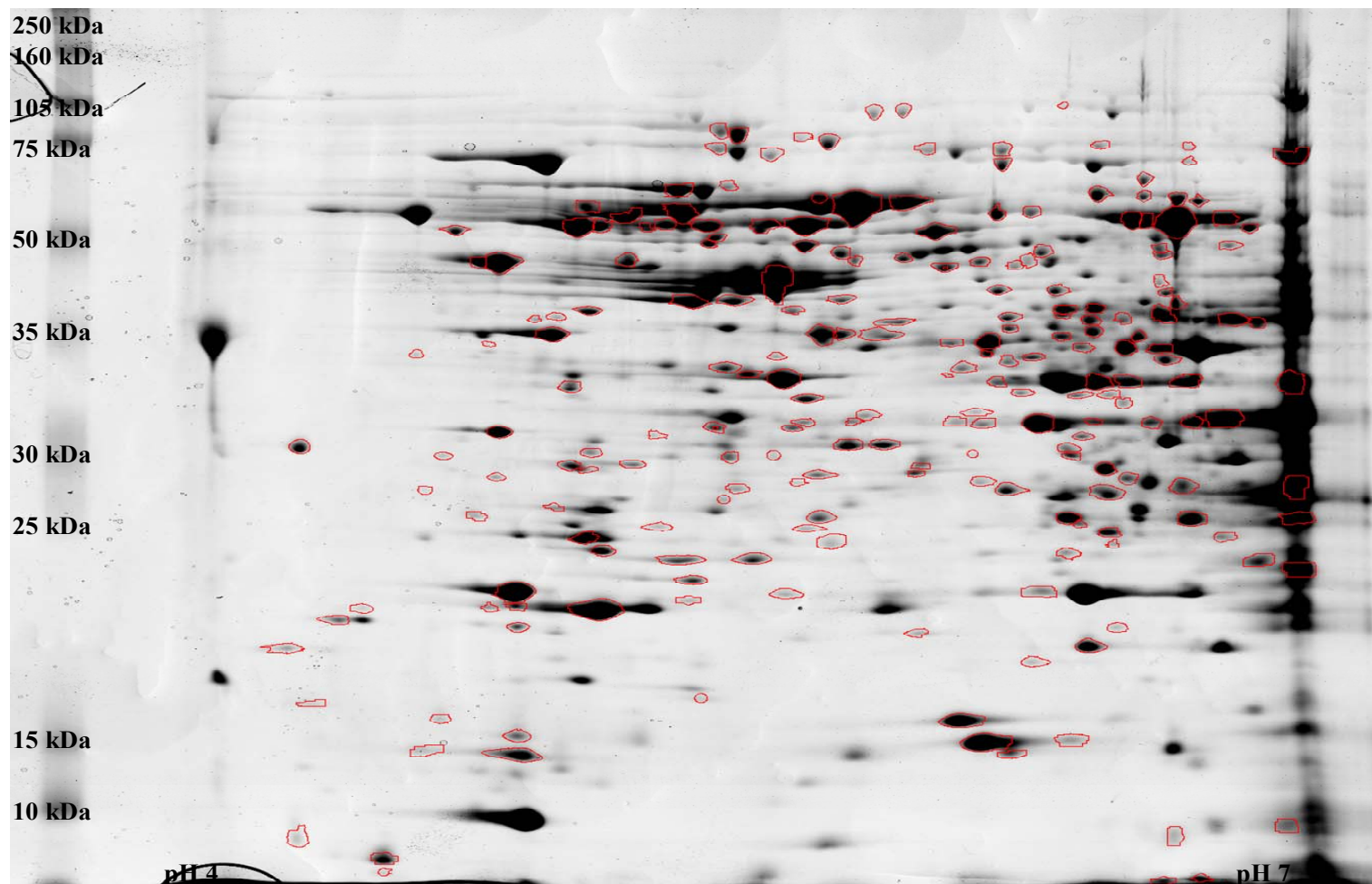


Figura 7: Análise por 2D-GE de 1 mg de proteínas de *G. diazotrophicus* cultivadas em meio LGI suplementado com 10mM de $(\text{NH}_4)_2\text{SO}_4$. À esquerda: padrão de peso molecular; em vermelho: *spots* identificados por espectrometria de massas.

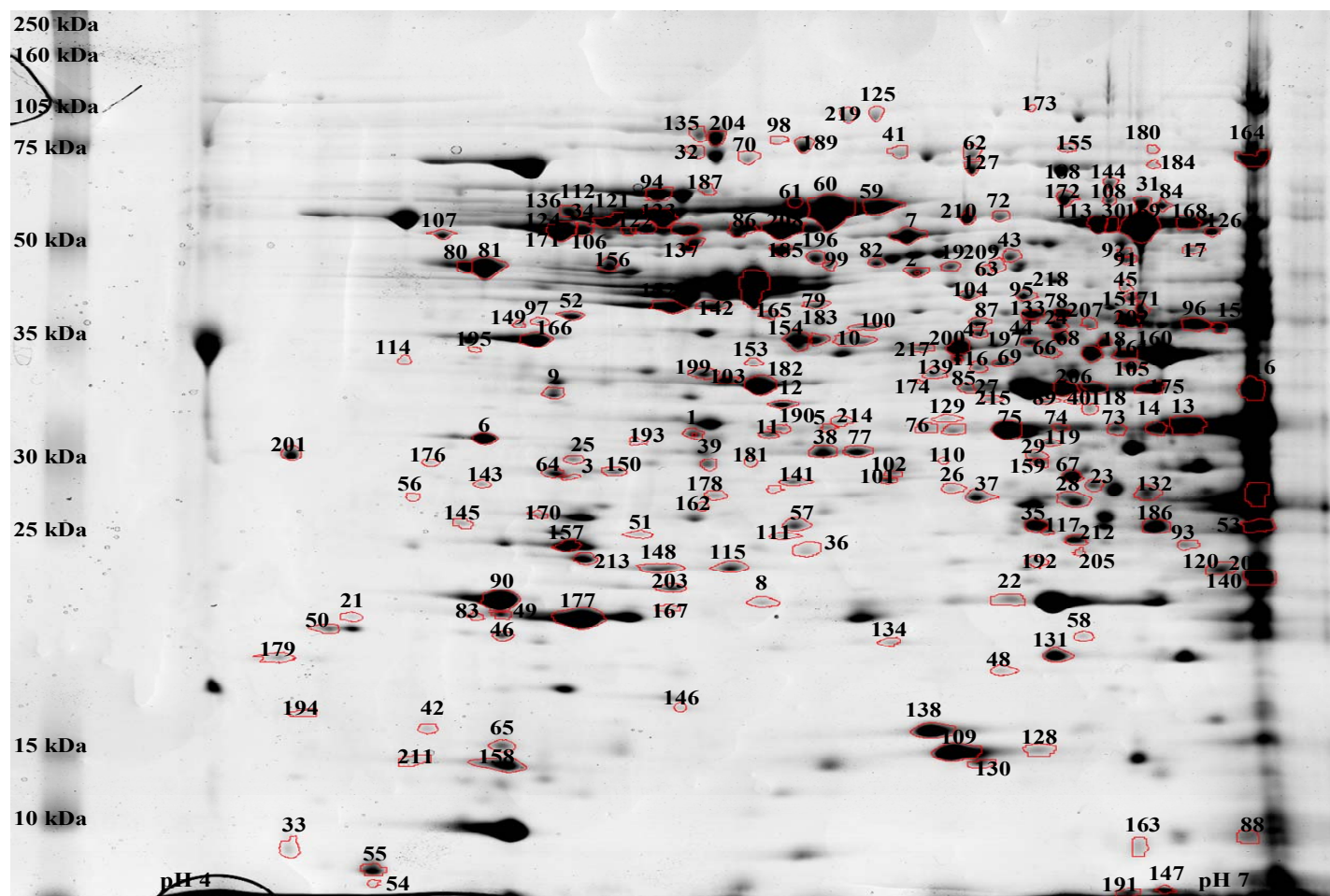


Figura 8: Análise por 2D-GE de 1mg de proteínas de *G. diazotrophicus* cultivadas em meio LGI suplementado com 10mM de $(\text{NH}_4)_2\text{SO}_4$. À esquerda: padrão de peso molecular; em vermelho: *spots* identificados por espectrometria de massas, numerados de acordo com a tabela 1 (anexo).

Todos os 354 *spots* visualizados no gel foram digeridos pela enzima tripsina e analisados por espectrometria de massa MALDI. Dos 354 *spots* iniciais, 208 foram identificados com confiança de $p < 0,05$, indicando 58% de sucesso na identificação (Fig. 8). Uma listagem completa de todas as proteínas e critérios de identificação estão relacionados na tabela 1 do anexo .

Alguns *spots* continham mais de uma proteína, portanto as 208 identificações iniciais derivaram de 196 *spots*, indicando co-migração de proteínas tanto na primeira quanto na segunda dimensão. Por exemplo, o *spot* c239 apresentou peptídeos que identificaram com confiança uma proteína fosfodiesterase carreadora de acil e uma proteína hipotética conservada. Ambas apresentam MM e pI teóricos parecidos: respectivamente 25 e 23 kDa e 6.1 e 5.8, tornando-as difíceis de serem separadas nas condições utilizadas neste trabalho.

As 208 identificações iniciais correspondem à 165 proteínas diferentes (números de acesso diferentes). Essa diferença demonstra a presença de um grande número de proteínas com modificações pós-traducionais, ou seja, isoformas de uma mesma proteína. Essas isoformas foram identificadas em *spots* com o mesmo valor de MM e uma pequena diferença no valor de pI. A presença de tais modificações é bem descrita em bactérias e em perfis bidimensionais de proteínas sempre são observados grupos de *spots* com mesmo valor de MM e pequenas diferenças de pI [120, 147].

4.1.3.2- Análise de MudPit

Assumindo que o genoma de *G. diazotrophicus* codifica em torno de 3000 proteínas (número de ORFs válidas em agosto de 2006 – Riogene), a identificação de 208 proteínas corresponde a apenas 7% da capacidade codificante prevista. Com o intuito de aumentar a taxa de proteínas identificadas em relação à proteínas previstas, decidimos utilizar uma nova estratégia, alterando o meio de cultivo, o método de lise celular, de separação das proteínas e análise para identificação.

Para tal finalidade, *G. diazotrophicus* foi crescida em meio DYGS e o extrato celular obtido foi submetido à eletroforese unidimensional (1D) em SDS-PAGE 12%. O gel 1D foi dividido em dez partes iguais ao longo do comprimento e cada uma das partes foi processada para espectrometria de massas, ou seja, foi descorada, desidratada e digerida *overnight* a 37°C com tripsina. Os peptídeos foram eluídos do gel com 50% de ACN e 0.1% de TFA e secos à metade do volume no *speed-vac*. Vinte microlitros desse eluído foi então injetado no Q-ToF Ultima API (Waters), desalinizado numa coluna C18 e, a seguir, foi separado cromatograficamente numa coluna C18. O eluído desta coluna foi diretamente analisado no espectrômetro de massas propriamente dito. Importante ressaltar que neste caso não obtemos uma informação de *fingerprint* de cada proteína, visto que cada amostra continha vários peptídeos de diversas proteínas. Por isto, neste caso, a obtenção de bons espectros de MSMS é fundamental para o seqüenciamento de peptídeos e indispensável à esta estratégia.

Das 10 partes do gel unidimensional contendo proteínas da amostra de *G. diazotrophicus* crescida em meio DYGS, nove (partes 2 a 10, fig. 9) foram analisadas conforme descrito anteriormente e resultaram num total de 13.872 espectros de MSMS. Este número tão alto deve-se ao fato que, por esta técnica, vários peptídeos de uma mesma proteína podem ser seqüenciados e de que certos peptídeos, principalmente os derivados de proteínas muito abundantes, podem ser analisados diversas vezes. Esses espectros foram utilizados para buscas nos bancos de dados do projeto genoma e culminaram com a identificação de 526 proteínas de *G. diazotrophicus* (fig.9).

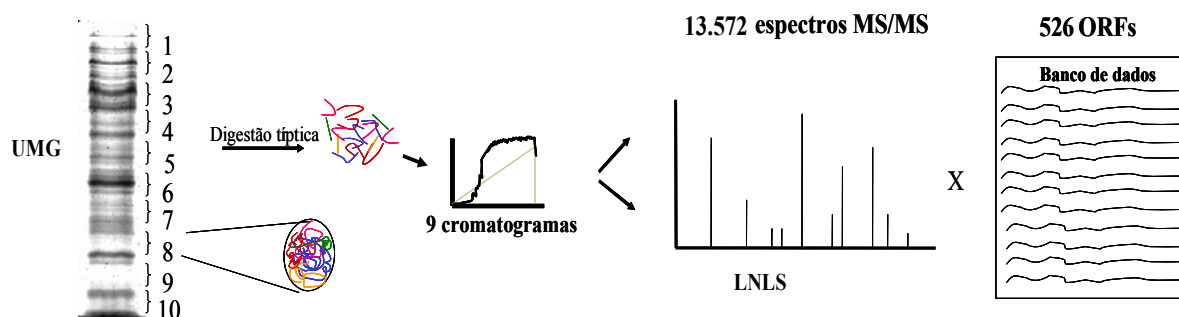


Figura 9: Esquema da análise de peptídeos por MudPit. Amostra de *G. diazotrophicus* crescida em meio DYGS foi analisada por SDS-PAGE e o gel resultante foi dividido longitudinalmente em 10 partes, digeridas com tripsina. 9 partes (partes 2 a 10) foram analisadas por LC/MSMS, resultando na identificação de 526 proteínas.

118 dessas proteínas já haviam sido identificadas por MALDI, de forma que considerando ambas estratégias (MALDI e MudPit) foram identificadas ao todo 573 proteínas diferentes (47 exclusivamente observadas em gel 2D MALDI, 118 por ambas estratégias e 408 exclusivamente por MudPit; figura 10).

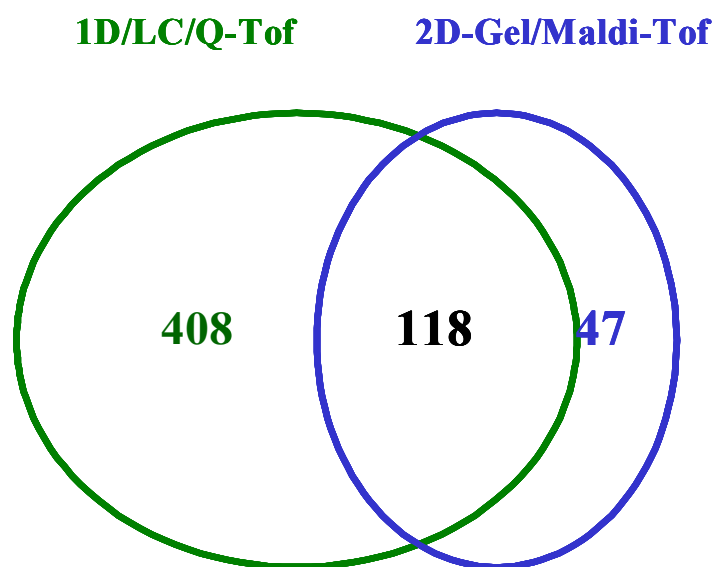


Figura 10: Diagrama esquemático mostrando as 573 proteínas de *G. diazotrophicus* identificadas: 118 identificadas tanto por análise em gel 2D e MALDI quanto por MudPit, 408 identificadas exclusivamente por MudPit e 47 visualizadas exclusivamente em gel 2D e identificadas por MALDI.

Com apenas as análises de 2D-GE e MALDI identificamos 7% da capacidade codificante de *G. diazotrophicus* (208 proteínas identificadas / 3000 ORFs preditas), enquanto que após as análises por MudPit cerca de 19% (573 proteínas / 3000 ORFs preditas) foi identificada. Dessa forma, a estratégia MudPit mostrou-se mais eficiente para identificação de proteínas em larga escala, embora algumas proteínas tenham sido identificadas apenas pela análise de 2D-GE. Este resultado não é tão surpreendente visto que os princípios dos métodos de separação das proteínas e peptídeos são diferentes. Pelo perfil cromatográfico derivado da análise das frações de gel 1D analisadas por LC/MS/MS (fig. 11), vê-se que diversos peptídeos foram co-eluídos da coluna. O espectrômetro de massas utilizado funciona alternando entre os modos de MS e MS/MS. Os peptídeos eluídos da coluna são detectados em modo MS. A cada vez que é detectado um pico intenso em MS, o segundo analisador de massa seleciona apenas a massa desse peptídeo e adquire seu espectro de MS/MS. Ao final do espectro de MS/MS o aparelho retorna ao modo MS, escaneando novos peptídeos eluídos da coluna. Assim, peptídeos que são co-eluídos da coluna não possuem grande chance de serem ambos identificados, pois apenas um espectro de MS/MS pode ser adquirido por vez. Para melhorar a separação dos peptídeos na coluna, pode-se ampliar o gradiente de eluição do peptídeos, permitindo que peptídeos com tempo de retenção muito próximos sejam adquiridos em tempos diferentes ou mesmo separá-los previamente em uma outra coluna cromatográfica, por exemplo uma troca iônica.

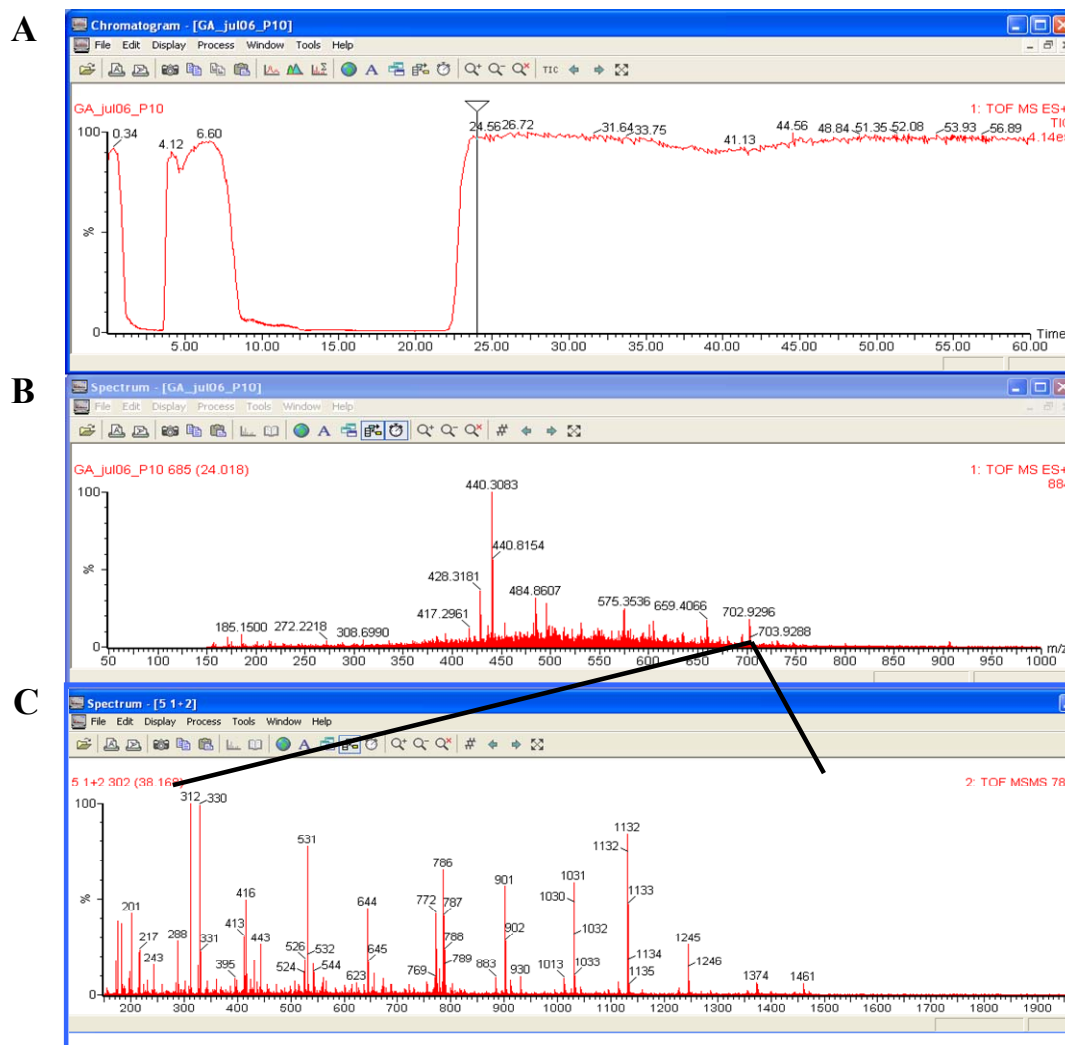


Figura 11: A) Perfil cromatográfico derivado da análise da fração 10 do gel 1D utilizado para análise por LC/MS/MS. O eixo das abcissas representa o tempo de eluição e o eixo das ordenadas indica a intensidade relativa do sinal, normalizada em 100% para o pico mais intenso. B) Espectro de massa do minuto 24,01 da corrida mostrada em A. C) Espectro de MS/MS do peptídeo de massa 787,2. Nos espectros de massa, o eixo das abcissas representa a razão massa/carga e o eixo das ordenadas indica a intensidade relativa do sinal, normalizada em 100% para o pico mais intenso.

4.1.4– Localização e categorização funcional das proteínas identificadas

4.1.4.1- Proteínas de membrana

A maioria das 573 proteínas identificadas neste estudo são solúveis. No entanto, algumas proteínas associadas à membrana também foram detectadas. Estas incluem duas lipoproteínas de membrana externa (*spots* c25, c313a), GTPase de membrana (*spot* c5), dipeptidase de membrana M19 (*spot* c156), subunidades da ATPase (*spot* c330) e proteínas de membrana prováveis (c109, c204 and c152). Proteínas associadas à membrana identificadas por 1D/LC/MSMS compreendem: proteína de secreção HlyD, proteína de exportação SecB, proteína de exportação de polissacarídeos, proteína ligadora de ATP do sistema de exportação de polissacarídeos capsulares CtrD, lipoproteína de membrana externa, proteína de membrana externa Omp16, proteína de membrana externa do sistema de efluxo e proteína de membrana YaeT. A identificação de algumas proteínas de membrana externa entre proteínas solúveis é usual, como já observado na análise de *V. cholerae* [91, 113] e outras bactérias.

A identificação de proteínas de membrana de *G. diazotrophicus* tem um interesse especial. Seu estilo de vida endofítico requer alterações do envelope celular como forma de adaptação aos diferentes ambientes encontrados. Além disso, já foi demonstrado que células de *G. diazotrophicus* apresentam características morfológicas distintas dependendo do meio extracelular e se estão ou não fixando nitrogênio [15]. Alterações ultraestruturais drásticas foram observadas em outros endófitos diazotróficos, incluindo alterações protéicas e remodelamento da membrana [148-150]. Assim, uma análise das proteínas de membrana de *G. diazotrophicus* em diferentes estados fisiológicos contribuiria para o melhor entendimento de suas funções na célula, assim como das vias reguladoras envolvidas.

4.1.4.2– Categorização das proteínas

As proteínas identificadas foram agrupadas de acordo com a classificação do Instituto Sanger para bactérias [151]. Na figura 12 é possível observar a distribuição das 573 proteínas

nessas categorias. Vale ressaltar as três categorias mais populosas: 220 proteínas envolvidas no metabolismo central e intermediário, 52 proteínas relacionadas com transferência de informação, ou seja envolvidas nos processos de replicação, transcrição, tradução e modificação de ácidos nucleicos e 42 proteínas do metabolismo de energia, incluindo cadeias de transporte de elétrons e glicólise.

Outras categorias destacadas foram: proteínas hipotéticas conservadas (40), de superfície (40) e hipotéticas (26, sendo 13 classificadas como proteínas desconhecidas e 13 como tendo alguma informação, mas com função incerta). Muito curioso foi a identificação de 34 proteínas reguladoras e 31 proteínas envolvidas na adaptação da bactéria a condições adversas.

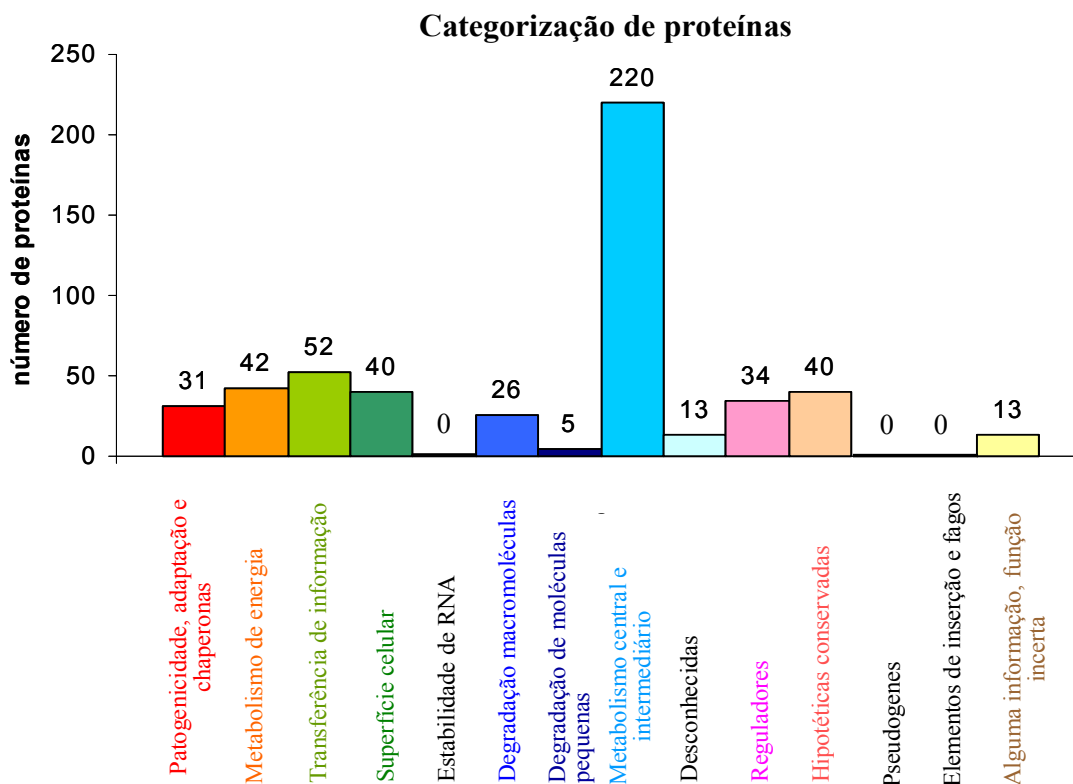


Figura 12: Categorização das proteínas identificadas de acordo com a definição do Instituto Sanger [151].

4.1.4.3– Vias metabólicas presentes em *G. diazotrophicus*

Dentre as 573 proteínas identificadas, foram assinalados *Enzyme Commission numbers* (EC) a 234. O EC indica a função de uma enzima e pode ser utilizado para identificar as vias metabólicas que a proteína participa, no banco de dados KEGG (Kyoto Encyclopedia of Genes and Genomes). Neste estudo, uma via metabólica foi considerada potencialmente ativa quando um mínimo de 3 enzimas da via tenham sido detectadas. Contudo, em diversos casos, como por exemplo na via glicolítica (fig. 13), muitas proteínas de uma mesma via foram encontradas, de forma que um quadro quase completo pode ser observado.

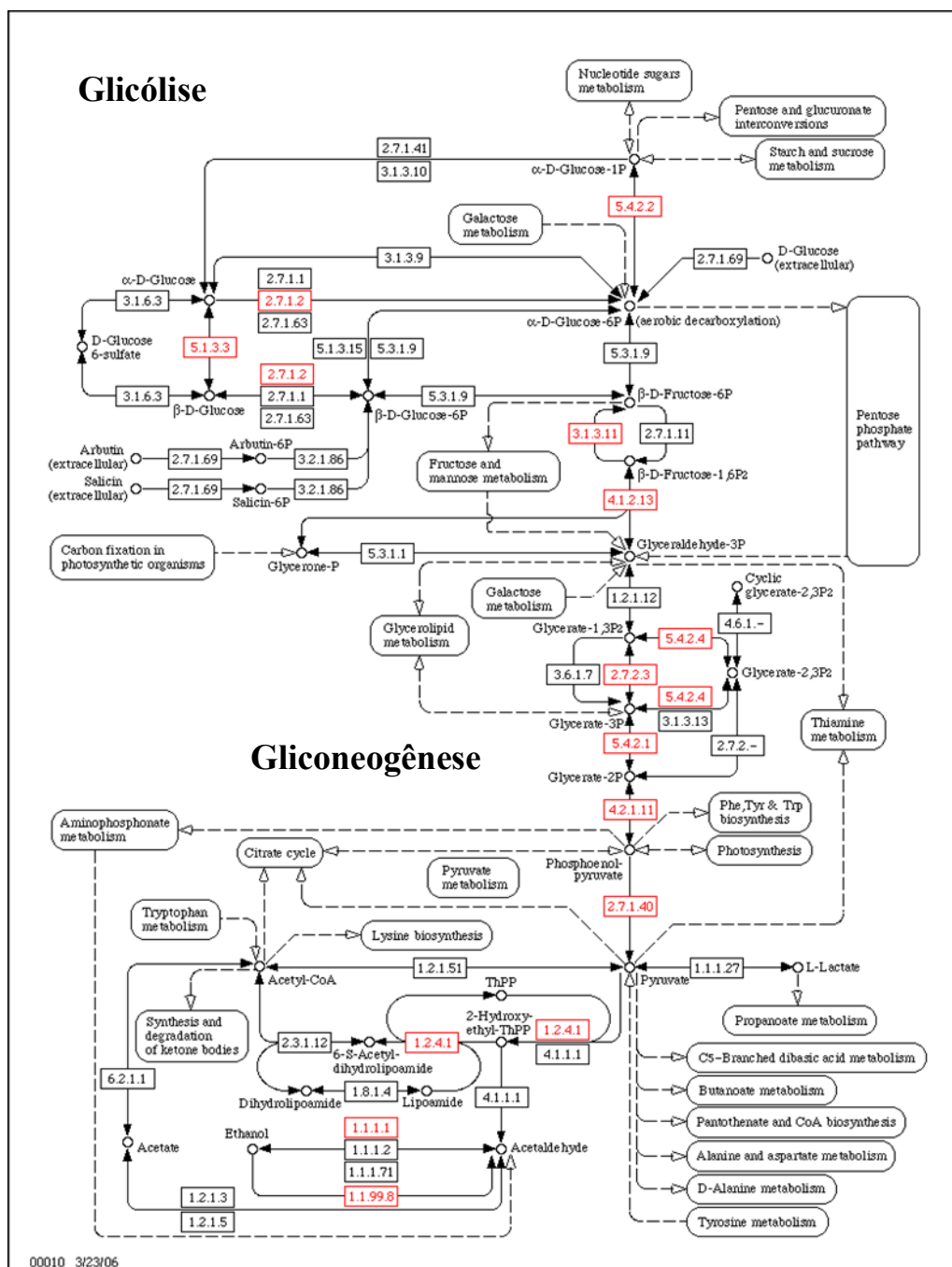


Figura 13: Esquema da via metabólica da glicólise e gliconeogênese gerado pelo KEGG. Nos quadrados estão representados os ECs das proteínas que participam desta via. Em vermelhos estão marcadas as proteínas identificadas no banco protéico de *G. diazotrophicus*.

Visto que tipicamente uma proteína pode participar de várias vias metabólicas, a identificação de uma ou duas proteínas em uma via pode ser mera casualidade e consequência da intrincada rede metabólica de uma célula. De acordo com este critério, 40 vias metabólicas foram representadas neste banco proteômico (fig. 14, tab. 2), permitindo uma visão global do metabolismo desta bactéria.

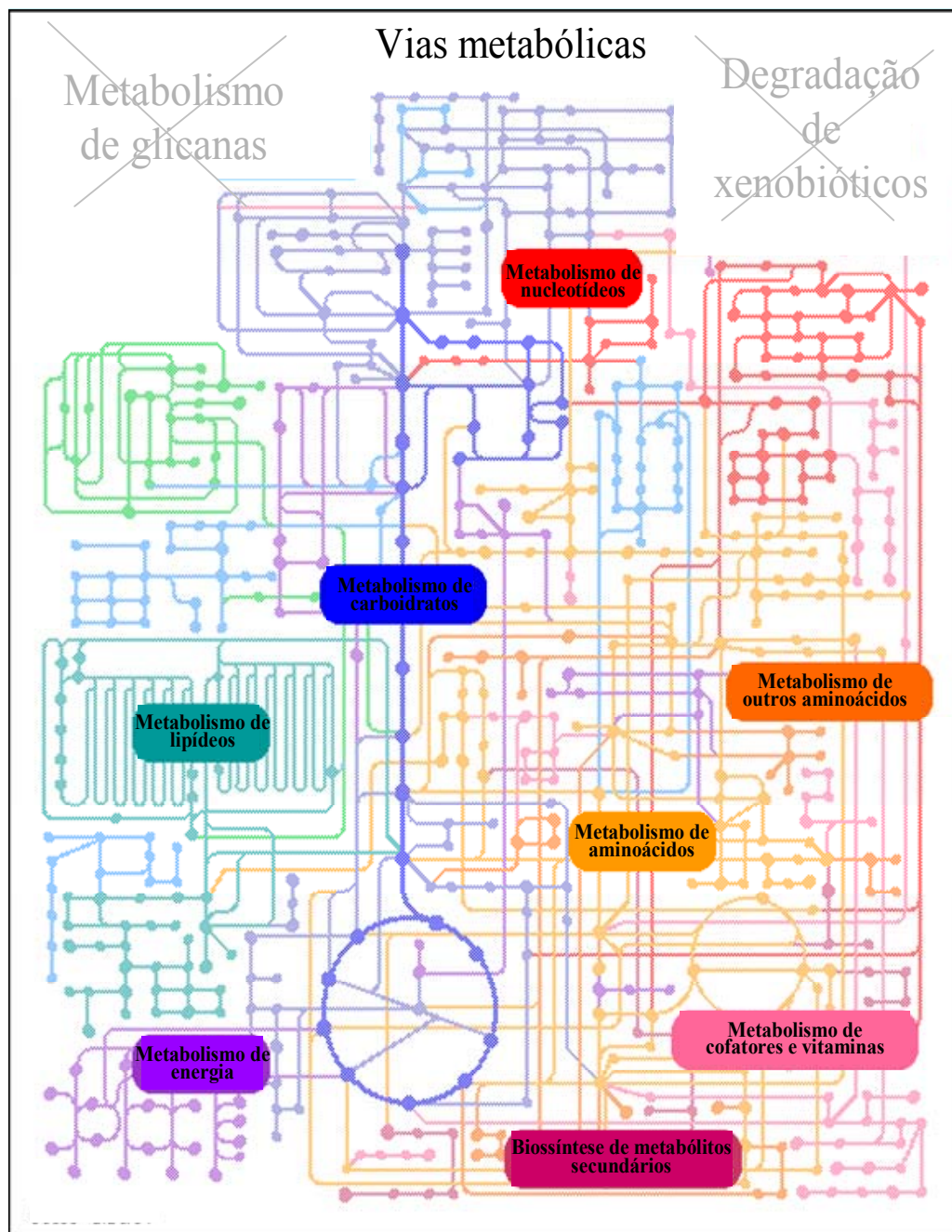


Figura 14: Esquema do metabolismo celular de *G. diazotrophicus*, mostrando as principais vias metabólicas identificadas no banco proteômico e como se relacionam numa rede integrada. Comparando este esquema obtido com o mapa do metabolismo celular de referência do banco de dados KEGG, verificamos que apenas as vias do metabolismo de glicanas e do metabolismo de xenobióticos não estão representadas com os dados deste trabalho.

Tabela 2: Vias metabólicas identificadas no banco proteômico de *G. diazotrophicus*.

	Vias metabólicas	Nº ECs
Metabolismo de Aminoácidos (16)	1- Metabolismo de glutamato	11
	2- Metabolismo de glicina, serina e treonina	9
	3- Metabolismo de metionina	9
	4- Biosíntese de fenilalanina, tirosina e triptofano	8
	5- Metabolismo de aspartato e asparagina	7
	6- Metabolismo de glutatona	7
	7- Biosíntese de leucina, valina e isoleucina	6
	8- Biosíntese de lisina	6
	9- Metabolismo de prolina e arginina	6
	10- Metabolismo de tirosina	6
	11- Ciclo da uréia e metabolismo de grupos amino	4
	12- Metabolismo de selenoaminoácidos	4
	13- Metabolismo de cisteína	3
	14- Metabolismo de fenilalanina	3
	15- Metabolismo de triptofano	3
	16- Metabolismo de beta-alanina	3
Metabolismo de Carboidratos (11)	17- Glicólise / Gliconeogênese	14
	18- Via das pentoses-fosfato	14
	19- Metabolismo do piruvato	8
	20- Ciclo do ácido cítrico	7
	21- Metabolismo de frutose e manose	5
	22- Metabolismo de butanoato	5
	23- Interconversões de pentose e glucoronato	4
	24- Metabolismo de galactose	4
	25- Metabolismo de amido e sacarose	4
	26- Metabolismo de nucleotídeos-açúcares	4
	27- Metabolismo de glioxilato e dicarboxilato	4

	28- Fixação de carbono	13
Metabolismo de energia (4)	29- Metabolismo de metano	7
	30- Metabolismo de nitrogênio	4
	31- Ciclo redutor de carboxilato	3
	32- Metabolismo de folato	4
Biosíntese de cofatores e vitaminas (2)	33- Metabolismo de porfirinas e clorofila	4
	34- Biosíntese de ácidos graxos	4
Metabolismo de lipídeos (2)	35- Biosíntese de esteróides	4
	36- Biosíntese de estreptomicina	6
Metabolismo de nucleotídeos (2)	37- Metabolismo de purinas	20
	38- Metabolismo de pirimidinas	16
Processamento de informação genética (1)	39- Biosíntese de aminoacil-tRNA	11
	40- Sistemas de dois componentes	5
Processamento de informação ambiental (1)		

Dentre as vias metabólicas identificadas, 16 são relacionadas ao metabolismo de aminoácidos, tais como de glutamato (11 proteínas), glicina, serina e treonina (9), metionina (9), fenilalanina, tirosina e tritofano (8), alanina e aspartato (7) e glutatona (7). Curiosamente, o metabolismo de selenoaminoácidos (4) e de beta-alanina (3) também estão representados.

Além disso, 11 vias do metabolismo de carboidratos foram caracterizadas (fig. 15). As mais representadas foram: glicólise e gliconeogênese (14 proteínas), via das pentoses fosfato (14), metabolismo de piruvato (8) e ciclo do ácido cítrico (7). Ainda, 13 proteínas da via de fixação de carbono, 7 do metabolismo de metano, 4 do nitrogênio e 3 do ciclo redutor de carboxilatos foram detectadas ativas, como parte do metabolismo de energia.

Em contraste, no metabolismo de nucleotídeos, representado em apenas duas vias o metabolismo de purinas e o de pirimidinas foram identificadas respectivamente 20 e 16 proteínas. Adicionalmente, 2 vias da biosíntese de cofatores, 2 do metabolismo de lipídeos, 1

da síntese de metabólitos secundários, 1 de processamento de informação genética e 1 de processamento de informação ambiental também foram encontradas.

Como esperado, as principais vias de manutenção do metabolismo celular, *housekeeping*, tais como glicólise e gliconeogênese, via das pentoses fosfato e síntese de diversos aminoácidos (alanina, aspartato, glicina, lisina, fenilalanina, serina, treonina e triptofano), nucleotídeos, sacarídeos, ácidos graxos, aminoacil-tRNAs e proteínas ribossomais estão ativas.

4.1.4.4- Proteínas envolvidas em transporte e outras proteínas de superfície celular

Das proteínas identificadas e que não possuem função enzimática, 21 são transportadoras ou parte de sistemas de transporte. Os sistemas de transporte do tipo ABC, em geral, incluem 1 proteína periplasmática de ligação ao substrato e 2 domínios de membrana interna, capazes de quebrar e ligar ATP. Vários componentes de sistemas ABC foram identificados, incluindo duas proteínas periplasmáticas e uma ligadora de ATP de sistemas de transporte ABC ainda não muito bem definidos. Além destas, proteínas ligadoras de ATP envolvidas nos sistemas de captação de açúcares e de tiamina foram detectadas. Ainda, SufC, uma ATPase que pode associar-se a outras proteínas Suf para a montagem de grupos Fe-S ou a proteínas de membrana para formar um complexo ABC de exportação [152] e uma proteína de ligação ao ATP do transportador ABC Uup [153] também foram encontradas.

Transportadores periplasmáticos de alta afinidade a D-ribose, D-xilose e fucose e os transportadores de fosfato PhoU e de biopolímeros ExbB foram detectados. Por fim, TolB, componente de um complexo de proteínas envolvido na importação e exportação [154] e TonB, o transdutor da energia próton-motiva do transporte ativo de ferro por sideróforos [154] também foram identificados. A disponibilidade de ferro livre é baixa em plantas e a captação de ferro é importante para relações bactéria-hospedeiro [155].

Proteínas envolvidas na montagem de estruturas da superfície celular ou exportação de moléculas também foram encontradas e incluem: SecB (chaperona que auxilia exportação de proteínas; [156]), HlyD (proteína do sistema de secreção do tipo I; [157]), proteína do sistema de secreção do tipo II/IV, CtrD, YaeT, uma UDP-N-acetilglucosaminilglicosaminidase (formação da parede celular, [158]), dois prováveis antígenos de superfície e diversas proteínas de membrana externa. Em *E. coli*, o gene *yaeT* codifica uma proteína essencial ao endereçamento e enovelamento de proteínas de membrana externa [159, 160].

4.1.4.5- Proteínas relacionadas à fixação de nitrogênio

A nitrogenase é um complexo protéico que cataliza a redução do dinitrogênio a amônia, acoplado à hidrólise de ATP [161]. Conforme descrito previamente, este complexo consiste de duas metaloproteínas, a Fe-proteína (NifH) e a MoFe-proteína (NifD e NifK) [162]. A expressão dos genes *nifHDK* é dependente do ativador transcricional NifA [35, 36]. Neste trabalho, identificamos as 3 proteínas da nitrogenase, NifH, NifD e NifK, assim como NifA e NtrX, que é um regulador de resposta do sistema de dois componentes envolvido na ativação de genes de assimilação de nitrogênio, tal como *glnA* [163].

4.1.4.6- Proteínas de adaptação e resposta de proteção

Entre as proteínas envolvidas na adaptação celular e respostas de proteção, foram encontradas 1 proteína relacionada com estresse em geral, 5 chaperonas, 3 proteínas do metabolismo de ferro, 2 proteínas antioxidantes, 2 de choque frio, uma proteína de tolerância a solventes orgânicos, a proteína de resistência a substâncias tóxicas EmrA, uma bacteriocina e a proteína induzida por paraquat (um composto organoclorado altamente tóxico [164]) PqiB.

As chaperonas identificadas são DnaK, DnaJ, HtpG, SurA e uma co-chaperona. As chaperonas moleculares compõem um sistema celular essencial, dependente de ATP, que possibilita o enovelamento e montagem de proteínas recém-sintetizadas, a translocação de proteínas não enoveladas pela membrana, assim como o re-enovelamento e degradação de

proteínas mal-enoveladas ou agregadas [165]. No mais, elas controlam vias de transdução de sinais por associações transientes com cinases e fatores de transcrição [166]. Uma das proteínas que provê assistência ao enovelamento protéico durante sua biossíntese é a chaperona bacteriana DnaK. A especificidade desses processos de enovelamento é determinada pela atividade de co-chaperonas, tais como DnaJ, que funciona juntamente com DnaK e GrpE em diversos processos celulares [167]. Em outro sistema, SurA, uma proteína requerida para a sobrevivência de *E. coli* durante a fase estacionária do crescimento, está envolvida com o enovelamento de proteínas [168]. Por outro lado, HtpG, componente do sistema bacteriano homólogo ao sistema mamífero HSP90, não é essencial em *E. coli* [169], embora crucial para resposta de estresse térmico em cianobactérias [170].

A maioria das chaperonas e proteínas associadas à elas que lidam com a alteração de conformação estrutural são expressas em resposta ao aumento da temperatura ambiente, determinado choque térmico (*heat shock*). Em contraste, após um decréscimo na temperatura, nomeado choque frio (*cold shock*), a resposta celular envolve predominantemente a ação de helicases, nucleases e componentes associados a ribossomos [171]. Dentre essas proteínas, a proteína de choque frio A com domínio DEAD e a proteína de choque frio CspE foram identificadas em *G. diazotrophicus*, indicando que essas proteínas devem desempenhar funções além de proteção ao choque frio, sendo ativadas em outras condições de estresse. Tais funções devem estar relacionadas com a estabilização de mRNA pela helicase DEAD e a condensação cromossômica pela proteína CspE, ligadora de ácidos nucleicos fita simples [172].

Duas proteínas envolvidas em mecanismos de proteção celular são a superóxido dismutase e peroxiredoxina, que participam das defesas da células contra espécies reativas de oxigênio (ROS). Superóxido dismutase converte superóxido a peróxido de hidrogênio e é uma das proteínas antioxidantes mais conhecidas [173]. Peroxiredoxinas, uma família de pequenas

proteínas antioxidantes que contêm resíduos de cisteína no sítio catalítico e usam a tioredoxina como doadora de elétrons, também convertem o peróxido e estão envolvidas na resposta celular aos ROS [174].

4.1.4.7- Proteínas reguladoras

Das proteínas identificadas, 34 possuem atividade reguladora prevista, contradizendo a idéia de que a análise proteômica é incapaz de detectar proteínas de baixa abundância. Essas proteínas possivelmente são reguladores globais, cujas funções no controle do metabolismo celular poderão ser avaliadas por mutações que inativem seus respectivos genes.

As proteínas reguladoras identificadas são: BolA, que em *E. coli* é essencial para alteração da morfologia celular e leva ao aumento da transcrição dos genes de síntese de parede celular *dacA* (PBP5), *dacC* (PBP6) e *ampC* [175]; proteína *B de utilização de N*, envolvida no processo de terminação da transcrição [176]; regulador transcricional ROS/MucR, que em *Rhizobium* regula diversas funções, incluindo a biossíntese dos exopolissacarídeos succinoglicano e galactoglicano [177]; regulador transcricional ChvI, parte de um sistema de dois componentes que regula a síntese de succinoglicanos [178]; e o regulador transcricional LysR.

Outros reguladores identificados são: proteína NusG, regulador transcricional da família IclR /regucalcina, regulador transcricional da família CarD, as proteínas cinases polifosfato cinase, inosina-guanosina cinase, ribocinase, ribulocinase, fosfoglicerato cinase, uridilato cinase, cinase PrkA e uma provável serino protease e as 3 proteínas da nitrogenase NifH, NifD and NifK, assim como NifA e NtrX.

4.1.5- Conclusões parciais

Este trabalho apresentou um primeiro e amplo conjunto de proteínas expressas por *G. diazotrophicus*. Destas, 208 proteínas foram identificadas a partir da análise por eletroforese bidimensional e 526 pela estratégia de MudPit. Ao todo, 573 proteínas diferentes foram

descritas e 40 vias metabólicas foram detectadas. A análise por MudPit possibilitou a identificação de um número maior de proteínas, no entanto, os valores de pI e MM da proteína madura só puderam ser obtidos através da eletroforese bidimensional.

IV.2- ANÁLISE DIFERENCIAL

4.2– ANÁLISE DIFERENCIAL

4.2.1– Cinética de crescimento

G. diazotrophicus foi cultivada em diferentes condições no Laboratório de Genética e Bioquímica da EMBRAPA Agrobiologia. As curvas de crescimento da bactéria crescida em meio LGI, pH inicial 5,5 e a 30°C em condições de limitação (L) e abundância (H) de compostos nitrogenados no meio foram feitas em duplicatas. Perfis semelhantes de crescimento foram obtidos para as culturas duplicatas, indicando a reprodutibilidade do experimento. Durante cada crescimento foram coletados de 5 a 15 amostras das culturas em tempos distintos. Uma vez que a fase *lag* das culturas duplicatas não tiveram a mesma duração, preferiu-se utilizar a mistura das amostras dos pontos da fase exponencial (E) como representativa da fase exponencial e a mistura dos pontos da fase estacionária (S) como representativa da fase estacionária. A cinética do crescimento de *G. diazotrophicus* PAL5 em L e H de compostos nitrogenados foi avaliada em meio LGI com pH inicial de 5,5, suplementado com 1 mM ou 10mM de $\text{NH}_4(\text{SO}_4)_2$ [1, 21] (fig. 15). Dados da literatura sugerem que 10mM de $\text{NH}_4(\text{SO}_4)_2$ inibe parcialmente a atividade da nitrogenase e portanto a fixação de nitrogênio por células de *G. diazotrophicus*, o que não ocorre com 1mM de $\text{NH}_4(\text{SO}_4)_2$ [21].

Os gráficos A, B e C da figura 15 mostram respectivamente a variação do número de células por ml de cultura, da concentração de $\text{NH}_4(\text{SO}_4)_2$ e do percentual de oxigênio dissolvido no meio de cultivo durante o crescimento. O acentuado consumo de $\text{NH}_4(\text{SO}_4)_2$ e O_2 entre a 15^a e 35^a horas, simultâneo ao grande aumento no número de células por ml, indicam uma alta taxa de multiplicação celular e caracterizam a fase E de crescimento. A estabilização do número de células e da concentração de $\text{NH}_4(\text{SO}_4)_2$ e o sutil aumento da concentração de O_2 , indicam o início da fase S. Por outro lado, a manutenção da concentração

de O_2 em valores baixos, mesmo na curva L (fig. 15C), é um fator adicional a associação desta condição de cultivo com o processo de fixação de nitrogênio.

Importante ressaltar que o pH do meio de cultura variou entre 5,5 (pH inicial) e 2,7 (dados não mostrados).

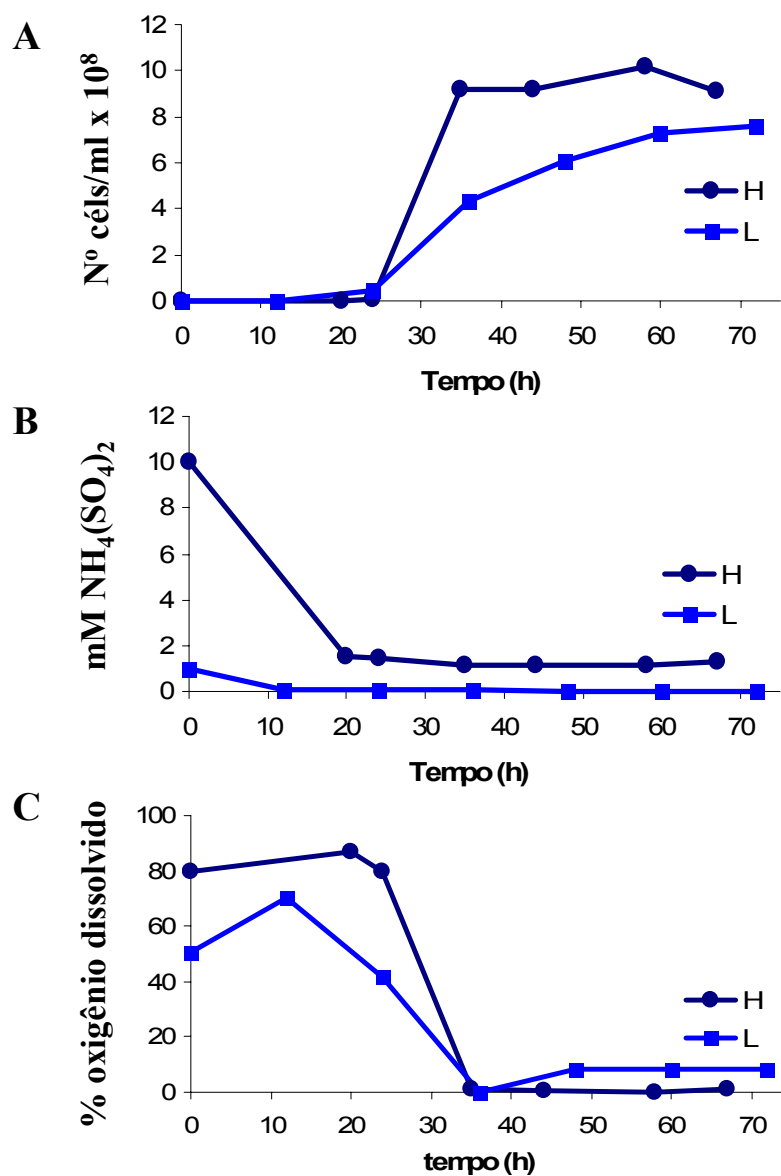


Figura 15: Cinética do crescimento de *G. diazotrophicus* em limitação (L) e abundância (H) de compostos nitrogenados. Em A: número de células x 10⁹ por ml de cultura; em B: concentração de $NH_4(SO_4)_2$; em C: % de O_2 dissolvido.

4.2.2– Análise proteômica: comparação de perfis protéicos obtidos por eletroforese bidimensional

Para cada condição experimental (EL, EH, SL, SH) foram preparados três lisados totais, sendo dois a partir de uma mesma amostra de uma curva de crescimento e outro a partir da amostra da outra curva. A comparação entre os géis bidimensionais dos 3 lisados de uma mesma condição experimental, permitiu a avaliação tanto da reprodutibilidade da amostra (comparação entre lisados de amostras de cultivos diferentes), quanto da reprodutibilidade da técnica 2D-GE (comparação entre os lisados da mesma amostra). A reprodutibilidade dos géis foi satisfatória em todos os casos.

A análise comparativa dos géis permitiu a obtenção de géis “master” EH, EL, SH and SL contendo respectivamente 381, 378, 375 e 372 *spots* (figs. 16-19). No total, 307 *spots* comuns a todos os géis, além dos *spots* específicos de cada condição (42 E + 32 S + 29 H + 28 L + 3 EH + 7 SH + 1 EL + 5 SL; fig. 20) foram visualizados distribuídos na faixa de pH 3 a 10, embora a maioria dos *spots* apresentaram-se na faixa 4 a 7. A grande maioria (~ 80%) das proteínas possuem MM variando entre 22 e 97 kDa, menos de 8 % com MM maior do que 97 kDa e cerca de 12 % menor do que 22 kDa.

Uma presença notável em todos os perfis 2D observados de *G. diazotrophicus* é um grande e alongado *spot* com pI em torno de 4 e MM de 45 kDa. Outra característica marcante é a presença de grupos de *spots*, possivelmente isoformas de proteínas. Muitas delas apresentaram pequenas alterações de pI, mas não de MM (figs. 16-19). Tais isoformas já foram identificadas em diversos organismos e representam um aspecto comum da regulação da função protéica [120, 129, 147].

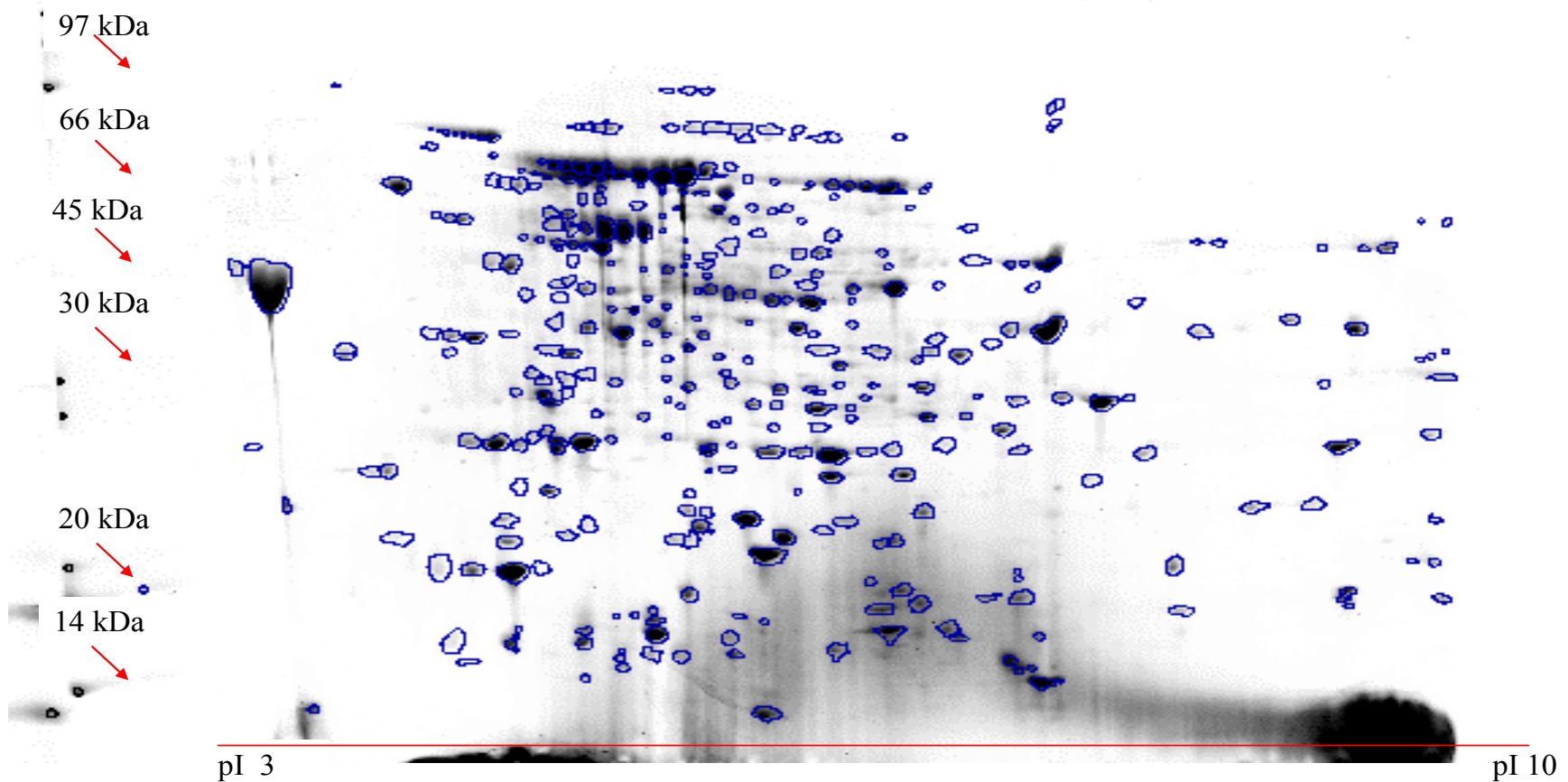


Figura 16: Sobreposição de seis géis virtuais representativos do padrão de expressão protéica de células de *G. diazotrophicus* em fase exponencial (E). Em azul, estão marcados os *spots* comuns às fases exponenciais das condições de limitação e abundância de N. As pontas das setas em vermelho indicam os marcadores de MM e a linha horizontal vermelha delimita a faixa de pI 3 a 10.

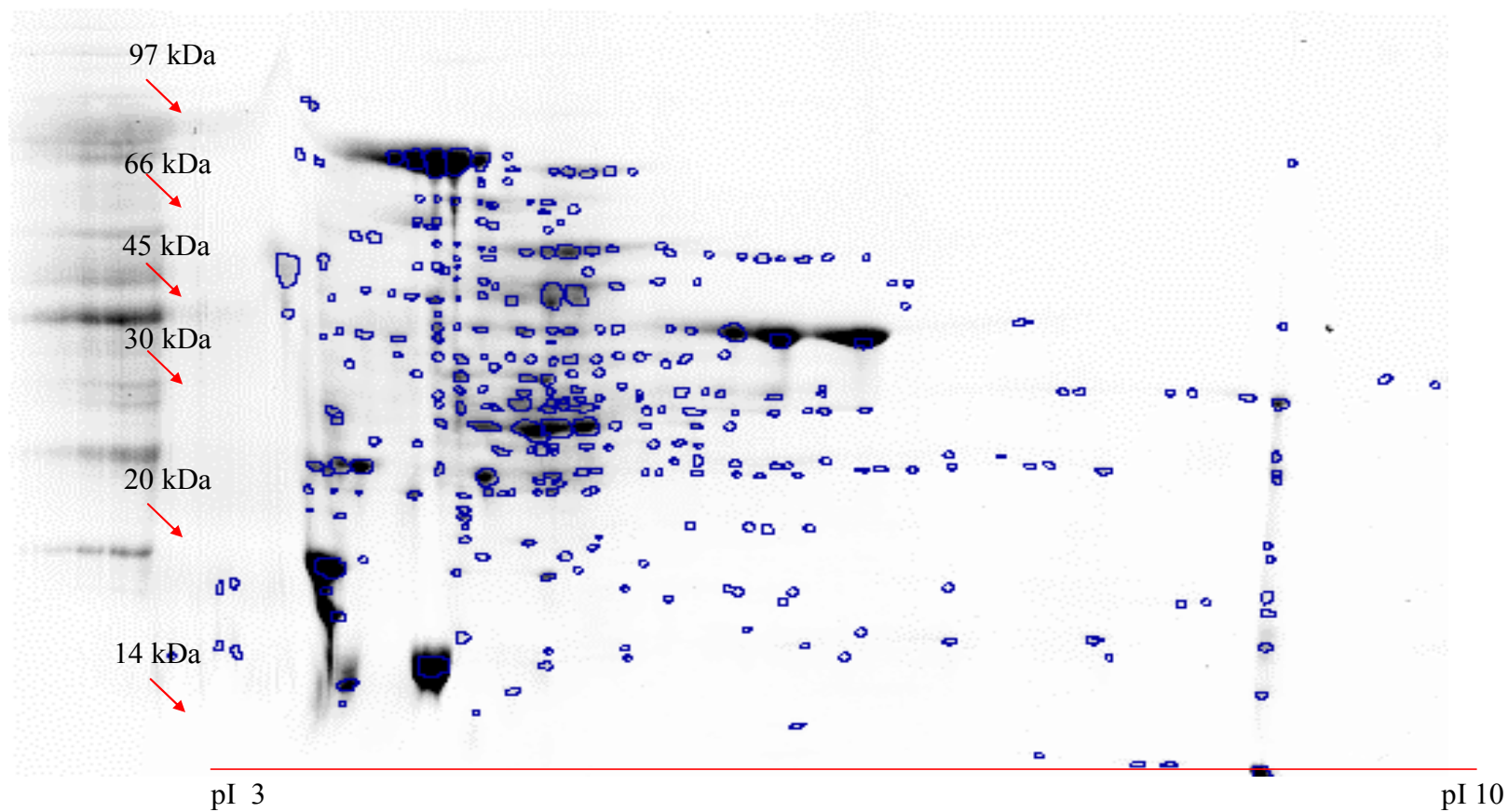


Figura 17: Sobreposição de seis géis virtuais representativos do padrão de expressão protéica de células de *G. diazotrophicus* em fase estacionária (S). Em azul, estão marcados os *spots* comuns às estacionárias das condições de limitação e abundância de N. As pontas das setas em vermelho indicam os marcadores de MM e a linha horizontal vermelha delimita a faixa de pI 3 a 10.

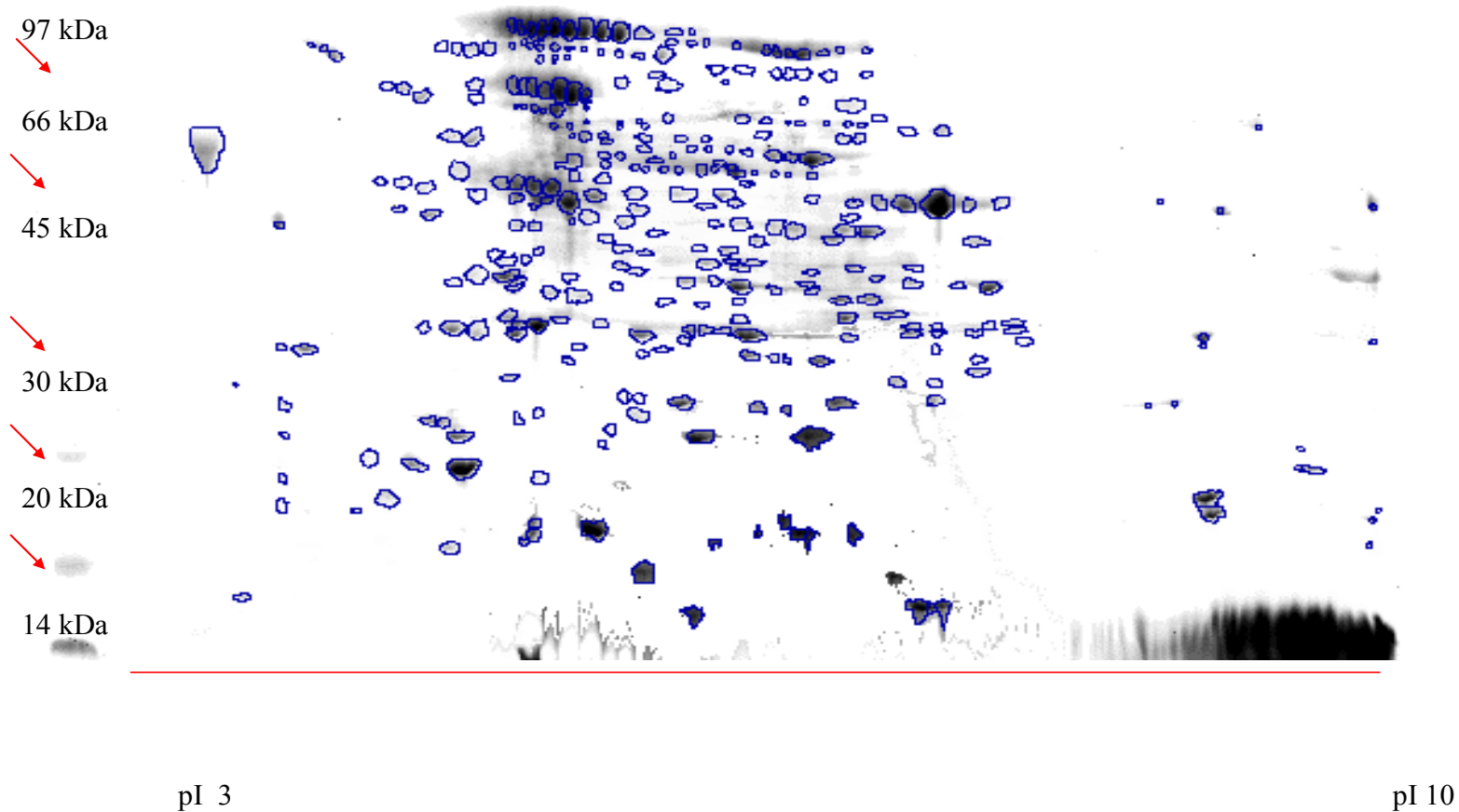
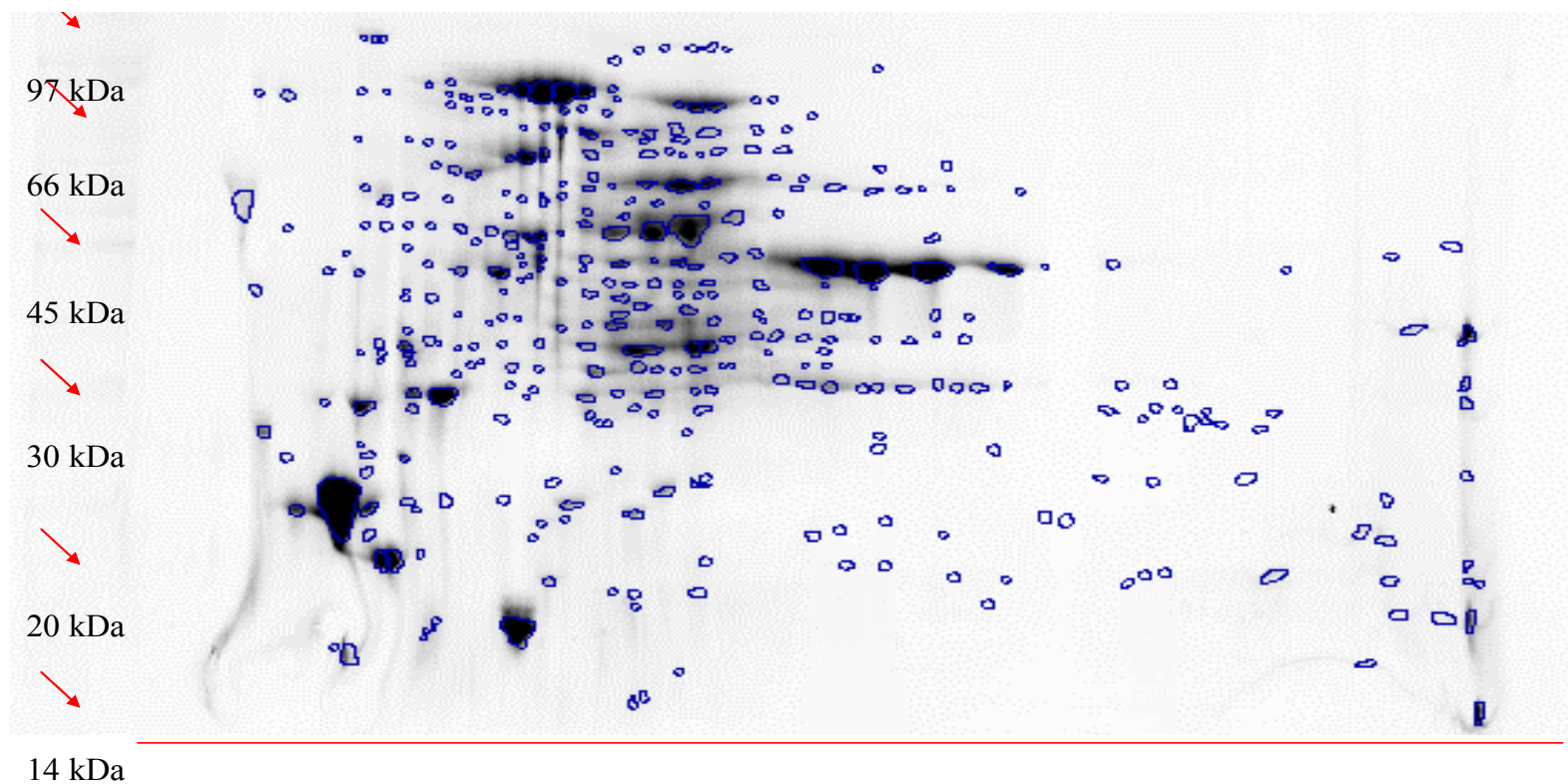


Figura 18: Sobreposição de seis géis virtuais representativos do padrão de expressão protéica de células de *G. diazotrophicus* em abundância (H) de N. Em azul, estão marcados os *spots* comuns às fases exponencial e estacionária da condição de abundância de N. As pontas das setas em vermelho indicam os marcadores de MM e a linha horizontal vermelha delimita a faixa de pI 3 a 10.



pI 3

pI 10

Figura 19: Sobreposição de seis géis virtuais representativos do padrão de expressão protéica de células de *G. diazotrophicus* em limitação (L) de N. Em azul, estão marcados os *spots* comuns às fases exponencial e estacionária da condição de limitação de N. As pontas das setas em vermelho indicam os marcadores de MM e a linha horizontal vermelha delimita a faixa de pI 3 a 10.

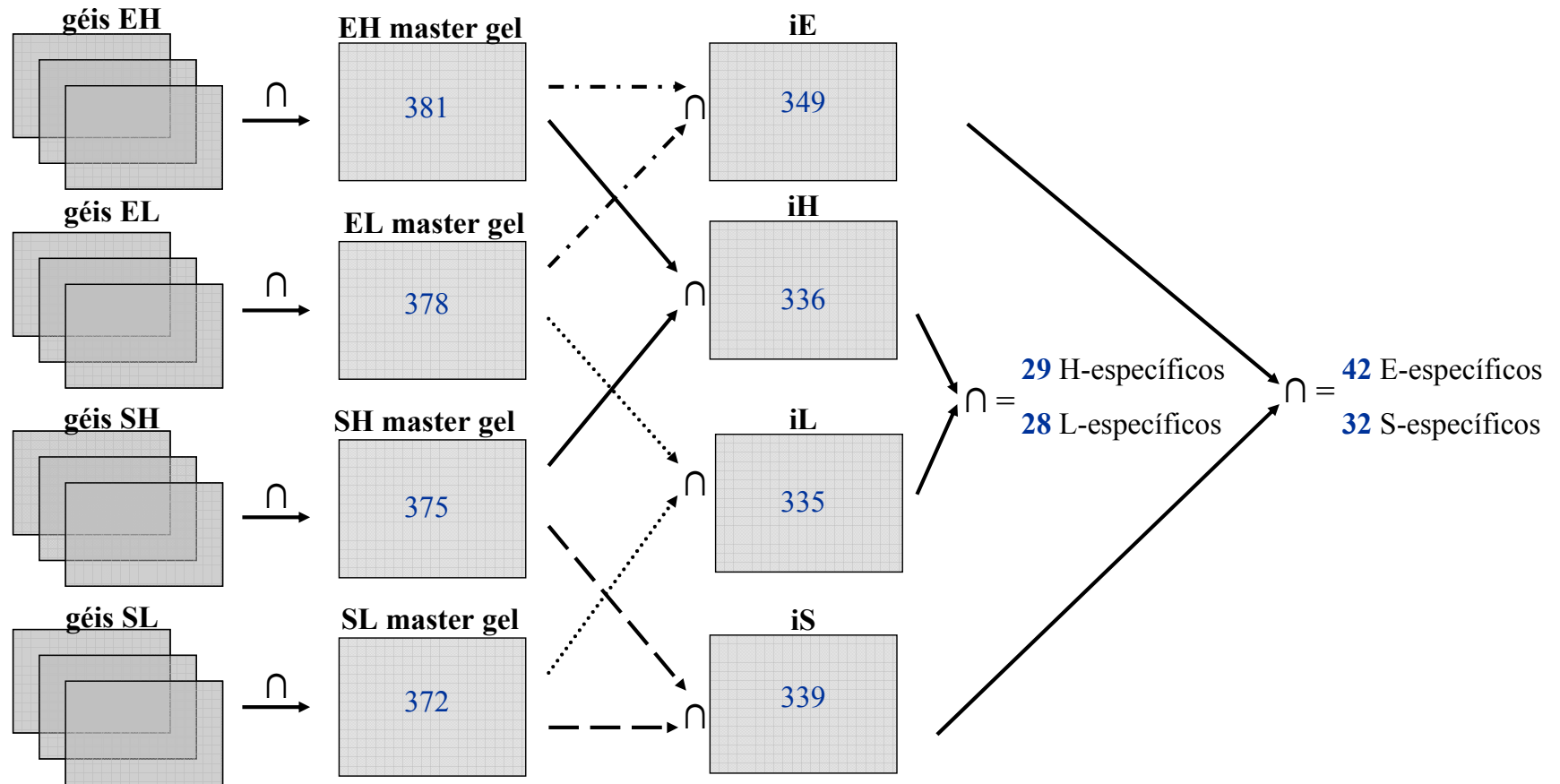


Figura 20: Design experimental da análise comparativa de géis, mostrando o número de *spots* analisados e diferencialmente expressos. Os géis membro foram sobrepostos e os *spots* comuns às triplicatas compuseram géis “master”. Interseções dos géis “master” foram obtidas e comparadas. *Spots* comuns a dois géis “master” foram considerados diferencialmente expressos quando seu volume relativo variou em mais de 3 vezes.

4.2.3- Identificação das proteínas expressas diferencialmente

Das 131 proteínas (42E + 32S + 29H + 28L = 131) expressas diferencialmente por *G. diazotrophicus* nas condições analisadas, 57 foram identificadas por espectrometria de massas. Destas, 39 foram identificadas pela análise perfil de peptídeos trípticos por MALDI-TOF. As proteínas não identificadas por esta técnica, foram submetidas a análise por MS/MS, gerando informações sobre sua sequência de aminoácidos.

Ao todo, foram identificadas: 40.5% (17 dos 42 *spots*) na fase E, 43.8% (14 dos 32 *spots*) da fase S, 78.6% (22 dos 28 *spots*) em L e 14% (4 dos 29 *spots*) de H (tabs. 3-6). Os demais *spots* diferencialmente expressos (74) estavam em quantidade insuficiente para serem analisados por MS ou MS/MS (possivelmente 28 *spots*) ou não puderam ser identificados nos bancos de dados.

Apenas *spots* com a mesma identificação no banco do NCBI (<http://www.ncbi.nlm.nih.gov>) e do Riogene foram considerados identificados. Proteínas identificadas apenas no banco do Riogene podem ser falso-positivos ou proteína únicas da espécie *G. diazotrophicus*. Proteínas exclusivamente encontradas no banco do NCBI podem ser produtos de ORFs não previstas, mal previstas, fora de *frame* ou divididas no banco do Riogene. Exemplos desses casos são mostrados na seção IV.3 desta dissertação.

As 57 proteínas identificadas são produtos de 48 genes. Por exemplo, as proteínas NodM e FixG foram ambas mais expressas na fase S de cultivo e na condição L, indicando que sejam especialmente requeridas em estágios tardios do crescimento, principalmente em condição de limitação de compostos nitrogenados no meio. (tabs. 4 e 5). A expressão diferencial das proteínas NodM e FixG nas condições acima não é uma surpresa, visto que ambas participam do processo de fixação de nitrogênio, que ocorre pelo aumento da atividade específica da nitrogenase durante a fase estacionária em algumas bactérias [179-182] e

provavelmente em *G. diazotrophicus*. Por outro lado, uma peptidase de sinal do tipo I e uma integrase são mais abundantes nas células da fase E e na condição H (tabs. 3 e 6).

Algumas proteínas nos géis foram encontradas em múltiplas isoformas, provavelmente como resultado de modificações pós-traducionais. Este foi o caso da lipoproteína associada a peptidoglicano e de uma proteína hipotética conservada, ambas com expressão aumentada na condição L.

4.2.4 – Categorização funcional das proteínas

As tabelas 3-6 apresentam os nomes e parâmetros de identificação das 57 proteínas caracterizadas neste trabalho, agrupadas por categorias de função celular definidas por Rilley [183].

A principal categoria induzida na fase E é formado por proteínas envolvidas nos processo de replicação, recombinação e reparo de DNA (tab. 3). Transportadores, por outro lado, compreendem a maioria das proteínas induzidas na fase S. Proteínas do metabolismo de energia, assim como do metabolismo de coenzimas foram identificadas tanto em células na fase E quanto em S (tabs. 3-4). Outras categorias induzidas na fase E são: tradução de sinais, motilidade e secreção celular, metabolismo de nucleotídeos, metabolismo de lipídeos e de aminoácidos. Outras classes de proteínas induzidas na fase S participam dos processos de fixação de nitrogênio, funções reguladoras, metabolismo de açúcares e tradução (tabs. 3-4).

Dentre as proteínas induzidas em L, os maiores grupos são de proteínas reguladoras, seguida por biogênese do envelope celular, fixação de nitrogênio e metabolismo de coenzimas. Outras proteínas induzidas em L atuam em processos de simbiose, metabolismo de energia, metabolismo secundário e tradução. Apenas 4 proteínas induzidas especificamente na condição H foram identificadas: 2 do metabolismo de aminoácidos, 1 envolvida em motilidade e secreção celular e a outra no metabolismo de DNA (tabs. 5-6)

Tabela 3: Proteínas induzidas em fase exponencial (E) do crescimento.

Spot	Proteína	Método	Cob	PM	Score
Replicação, recombinação e reparo de DNA					
1	ApaLI methyltransferase	MS	21	5	1,46E+07
10	Integrase	MSMS	65	25	1,33E+07
13	Transposase 4 ISRM30	MSMS	22	12	1,56E+07
14	Resolvase	MS	42	9	4,10E+04
17	Transposase	MS	53	13	1,07E+07
Metabolismo de Energia					
2	ATP synthase delta chain	MS	77	12	7,19E+04
11	Malate synthase	MS	40	9	1,01E+07
Metabolismo de Coenzimas					
4	Biotin synthase	MS	34	14	2,00E+07
9	Molybdenum cofactor biosynthesis protein A	MSMS	35	11	1,79E+07
Metabolismo de aminoácidos					
12	Proline dipeptidase	MSMS	46	16	1,44E+07
Metabolismo de purinas, pirimidinas, nucleotídeos e nucleosídeos					
6	Nucleoside diphosphate kinase	MS	64	20	1,84E+07
Metabolismo de lipídeos					
3	Beta-ketoacyl CoA thiolase	MSMS	44	18	3,84E+07
Biogênese de envelope celular					
8	Lipopolysaccharide biosynthesis protein	MSMS	46	25	1,96E+07
Motilidade e secreção celular					
16	Signal peptidase I	MS	38	11	1,33E+07
Tradução					
5	Deoxyhypusine synthase	MS	53	20	1,12E+07
Sinalização celular					
15	Sensor histidine kinase / response regulator	MS	41	16	3,17E+07
Outra					
7	Hypothetical	MS	60	16	1,63E+07

Cob: % da sequência da proteína coberta pelos peptídeos; PM: número de peptídeos identificados. Em negrito: proteínas cuja expressão variou em mais de 5 vezes.

Tabela 4: Proteínas induzidas em fase estacionária (S) do crescimento.

Spot	Proteína	Método	Cob	PM	Score
Transportadores					
18	ABC transporter, binding protein	MS	36	19	1,07E+07
24	Multiphosphoryl transfer protein	MS	37	18	1,49E+07
29	Beta 1,2 glucan export ATP-binding protein	MSMS	31	14	1,89E+07
20	ATP-binding protein of ribose ABC transporter	MS	40	18	3,50E+07
Outra / transferases					
28	Probable methyl transferase	MS	34	9	2,48E+07
22	Glucosyltransferase protein	MSMS	54	10	3,76E+07
19	Amidophosphoribosyl transferase-like protein	MS	46	16	1,02E+07
26	N-carbamyl-L-amino acid amidohydrolase	MS	56	24	3,35E+07
Metabolismo de Coenzimas					
23	Mg-protoporphyrin IX chelatase	MS	50	20	7,49E+07
27	Precorrin-6 γ methylase	MS	36	12	1,74E+07
Metabolismo de Carboidratos					
31	Ribulose biphosphate carboxylaseoxygenase, large subunit	MS	42	16	2,08E+08
Metabolismo de energia					
25	NADH-ubiquinone dehydrogenase chain E 1	MS	32	10	4,79E+07
Fixação de nitrogênio					
21	FixG protein	MS	45	20	2,50E+07
Tradução					
30	Putative leucyl-tRNA synthetase	MS	27	20	4,94E+07

Cob: % da sequência da proteína coberta pelos peptídeos; PM: número de peptídeos identificados. Em negrito: proteínas cuja expressão variou em mais de 5 vezes.

Tabela 5: Proteínas induzidas em limitação (L) de NH₄(SO₄)₂.

Spot	Proteína	Método	Cob	PM	Score
Funções Regulatórias					
32	Anaerobic regulator	MS	65	18	1,28E+07
40	Conserved hypothetical	MS	44	4	6,45E+04
41	Conserved hypothetical	MS	52	7	9,95E+04
42	Conserved hypothetical	MS	48	5	7,11E+04
49	ExsB protein	MSMS	33	5	1,10E+07
53	NifR3	MSMS	39	10	2,01E+09
36	Transcription activator MatR	MS	46	9	1,07E+07
Biogênese do envelope celular					
35	Penicillin-binding protein 2	MS	41	16	2,92E+07
45	Peptidoglycan associated lipoprotein	MSMS	46	7	1,28E+04
46	Peptidoglycan associated lipoprotein	MSMS	41	5	1,02E+04
47	Peptidoglycan associated lipoprotein	MSMS	37	4	1,99E+04
48	Peptidoglycan associated lipoprotein	MSMS	46	6	1,79E+04
Fixação de nitrogênio					
38	FixG protein	MS	45	20	2,50E+07
44	Nitrogen fixation protein W	MSMS	37	6	1,96E+04
50	NifB protein N-terminal	MSMS	45	10	1,87E+04
Metabolismo de coenzimas					
33	HypA protein	MS	55	9	7,46E+07
34	HypX protein	MS	39	19	1,93E+08
Simbiose					
39	Nodulation protein nodC	MS	30	21	1,39E+07
52	Nodulation protein M	MSMS	30	14	1,10E+04
Metabolismo de energia					
43	Dihydrolipoamide acetyltransferase component of pyruvate dehydrog	MS	51	7	3,55E+04
Metabólitos secundários					
37	Dienelactone hydrolase	MS	42	13	4,38E+07
Tradução					
51	Ribosomal protein L25	MSMS	39	11	2,56E+04

Cob: % da sequência da proteína coberta pelos peptídeos; PM: número de peptídeos identificados. Em negrito: proteínas cuja expressão variou em mais de 5 vezes.

Tabela 6: Proteínas induzidas em abundância (H) de $\text{NH}_4(\text{SO}_4)_2$.

Spot	Proteína	Método	Cob	PM	Score
Metabolismo de aminoácidos					
55	Aminopeptidase P	MS	22	11	1,67E+07
57	Urease D protein	MS	30	6	3,30E+04
Motilidade e secreção celular					
56	Signal Peptidase I	MS	38	11	1,33E+07
Replicação, recombinação e reparo de DNA					
54	Integrase	MS	45	25	6,03E+06

Cob: % da sequência da proteína coberta pelos peptídeos; PM: número de peptídeos identificados. Em negrito: proteínas cuja expressão variou em mais de 5 vezes.

4.2.5- Análise das proteínas induzidas na fase exponencial (E)

A maioria das proteínas diferencialmente expressas na fase E (fig. 21, tab. 3) estão envolvidas no metabolismo de DNA e processos de produção e conversão de energia, em acordo com a alta taxa de multiplicação celular e o alto requerimento de energia característicos desta fase.

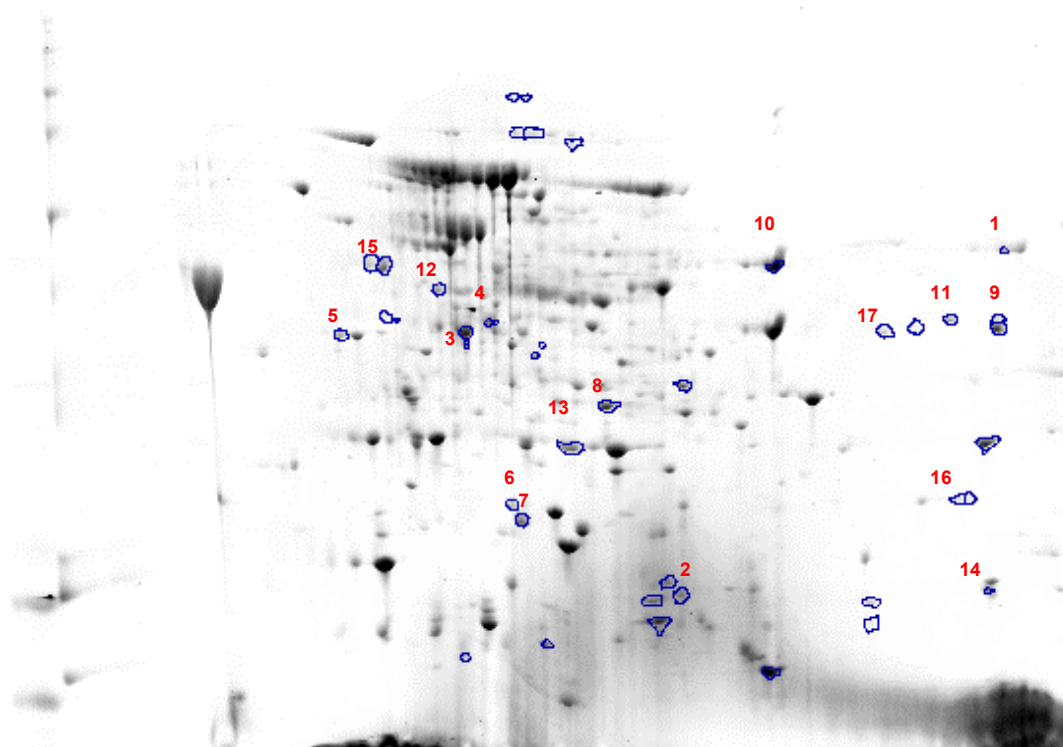


Figura 21: Perfil de 2D-GE de *G. diazotrophicus* em fase exponencial (E) de crescimento. Em azul estão destacados os *spots* induzidos nesta fase e os números em vermelho indicam as proteínas identificadas (tab. 3).

Proteínas com funções reguladoras também foram detectadas, tal como uma metiltransferase específica de citosina (*spot 1*, fig. 21). Esta proteína metila especificamente citosinas no DNA e provavelmente está envolvida na modificação do DNA celular e proteção à material genético externo [184]. Outra proteína com potencial função reguladora é uma histidina cinase de um canal de K^+ osmosensível (KdpD, *spot 15*, tab. 3). Em *E. coli* e outras bactérias, KdpD/KdpE são respectivamente uma histidina cinase sensora e um regulador de resposta do sistema que regula a expressão do operon *kdpFABC*, que codifica um

transportador de K^+ de alta afinidade [185, 186]. O influxo de K^+ na células de *G. diazotrophicus* na fase E pode ser uma estratégia para manter a homeostasia do pH citoplasmático. Neste estudo, as células foram cultivadas sem controle do pH do meio de cultivo. O pH inicial do meio foi 5,5, mas durante o crescimento o pH externo (pH_o) caiu a valores muito baixos (~2.7). Na literatura são descritos muitos casos de bactérias que tendem a manter o pH intracelular (pH_i) relativamente constante, independente de flutuações no pH_o [187]. Além disto, em muitos organismos, incluindo *E. coli*, o pH citoplasmático é aumentado pela extrusão de H^+ durante o influxo de K^+ [188]. Curiosamente, os sistema de transporte Kdp de *E. coli* é eletrogênico, levando à despolarização da membrana associada ao bombeamento de H^+ para o meio externo e aumento do pH_i [189-191].

Entre as proteínas envolvidas no metabolismo de energia, a subunidade β da ATP sintase (AtpD - *spot2*, tab. 3) foi identificada. ATP sintase é uma enzima que catalisa a síntese respiratória de ATP acoplada a um fluxo exergônico de prótons pela membrana plasmática da bactéria [192]. Tentativas de mutar o gene *atpD* de *R. capsulatus* e *Mycobacterium smegmatis* não tiveram sucesso, sugerindo que a ATP sintase seja essencial para estes organismos [192-194]. A expressão de uma AtpD por *G. diazotrophicus* em condições de pH não controlado é uma descoberta interessante e pode fazer parte da mesma estratégia que envolve o sistema Kdp, mencionado acima, para manter a homeostasia do pH citoplasmático. *G. diazotrophicus* é uma bactéria tolerante ao ácido e sua habilidade de sobreviver em ambientes ácidos pode ser aumentada por mecanismos que mantêm o pH_i. Em diversas bactérias tolerantes a ácidos, o pH da cultura cai durante o crescimento e o ATP citoplasmático é rapidamente perdido. No entanto, se açúcares forem adicionados ao meio, o nível de ATP intracelular aumenta. Aparentemente, parte desse ATP é oriundo da ATP sintase ativada pelo baixo pH, sugerindo que as células possuam duas fontes de síntese de ATP em baixo pH [195]. O ATP produzido, por sua vez, deve ser utilizado como substrato para F-

ATPases que realizam a extrusão de prótons, contribuindo dessa forma a manutenção do pH_i, requerida para o crescimento [196].

Outras duas proteínas induzidas na fase E são o cofator A de molibdênio (*spot* 9, tab. 3) e a biotina sintase (*spot* 4, tab. 3). Molibdênio (Mo) apresenta atividade biológica quando complexado a um cofator, formando o cofator Moco [197]. Enzimas que contêm Moco catalisam importantes reações redox nos ciclos do carbono, enxofre e nitrogênio [198]. A biotina sintase (BioB; *spot* 4, tab. 3), por outro lado, catalisa o último passo na via de síntese de biotina (vitamina H). Em diversas bactérias, a biotina funciona como cofator de um pequeno número de enzimas, mas também apresentam funções não catalíticas, regulando e modulando a expressão gênica [199-201]. Na bactéria diazotrófica *S. meliloti*, diversos genes da via de síntese de biotina estão ausentes. Visto que a maioria dos organismos do solo têm a habilidade de sintetizar biotina [202, 203], a presença de uma via incompleta é curiosa. Como a biotina não está disponível na maioria dos ambientes naturais, ter a via de síntese de biotina pode significar uma vantagem seletiva. Muito pouco é sabido sobre a regulação dos genes desta via em *G. diazotrophicus*, contudo, seu crescimento é estimulado pela adição de biotina ao meio de cultivo [21, 204]. Assim, a síntese de BioB por *G. diazotrophicus* na fase E pode ser significativa e com a finalização do projeto genoma, mais informações sobre esta via metabólica poderão ser obtidas.

A malato sintase (*spot* 11, tab. 3) é uma enzima do ciclo de glioxilato (GLOX), que cataliza a condensação de glioxilato com acetil-CoA para formar malato, numa via alternativa ao ciclo do ácido tricarboxílico (TCA) [205]. As enzimas chave deste ciclo são a isocitrato liase e malato sintase, que ocorrem amplamente em bactérias, fungos, algas e plantas, quando em meios deficientes de açúcares [206]. Interessante é que em *Bradyrhizobium japonicum*, bactéria dizotrófica da soja, 50% do acetil-CoA que entra no ciclo do ácido tricarboxílico é utilizada pela malato sintase [207]. Além disso, a atividade específica da isocitrato

desidrogenase, outra enzima do TCA, diminui com o aumento da atividade da nitrogenase durante a simbiose [208], sugerindo baixa atividade do TCA durante o processo de fixação de nitrogênio. Aparentemente, a malato sintase é expressa constitutivamente em *B. japonicum* [206, 209], em outros *rhizobium* [210] e *Pseudomonas ovalis* [211]. Como o glioxilato é um intermediário de outras vias metabólicas além do GLOX [206, 212], é possível que a malato desidrogenase seja expressa *rhizobia*, *Pseudomonas* [213] e mesmo em *G. diazotrophicus* com funções não relacionadas ao metabolismo do acetato. A expressão desta enzima por *G. diazotrophicus* requer estudos funcionais específicos.

A β -cetoacil-CoA tiolase (*spot3*, tab. 3), produto do gene *pcaF*, também foi induzida especificamente na fase E de crescimento. PcaF atua na fase final da degradação de muitos compostos aromáticos, que são convertidos a catecol e protocatecato para produzir compostos que, por sua vez, entram no TCA e em última análise geram energia [214]. Esta via está presente em diversos microorganismos terrestres, possibilitando a utilização de compostos aromáticos produzidos em abundância pelas plantas. Curiosamente, diversos estudos mostraram que compostos aromáticos, tais como o benzoato e o 4-hidroxibenzoato são fortes quimioatratantes para certas bactérias Gram-negativas, incluindo *Agrobacterium*, *Azospirillum* e *Rhizobia*, espécies que colonizam plantas [215, 216]. A expressão de PcaF na fase E é um indicativo que *G. diazotrophicus* pode utilizar compostos aromáticos como fonte de carbono e que a quimiotaxia a esses compostos pode estar envolvida nas interações de *G. diazotrophicus* com seus hospedeiros.

A nucleosídeo difosfato cinase (NDK; *spot6*, tab. 3) é requerida para a síntese de nucleosídeos trifosfato que não o ATP [217]. Esta enzima provê NTPs para a síntese de ácidos nucleicos, CTP para síntese de lipídeos, UTP para síntese de polissacarídeos e GTP para o alongamento de proteínas, transdução de sinais, síntese protéica e metabolismo de

energia [217]. Assim, NDK funciona no crescimento bacteriano, virulência, sinalização celular e síntese de polissacarídeos [218].

Quatro proteínas específicas da fase E estão envolvidas em processo de recombinação não homóloga do DNA, que gera rearranjos no DNA, tais como inversões, excisões e co-integração [219]. *Spot* 10 é aparentemente um membro da família de integrases de fagos e *spot* 14 é uma resolvase. As outras duas, *spots* 13 e 17, são transposases envolvidas na recombinação por transposição. A maioria das recombinases específicas incluem integrases, resolvases, invertases e transposases, codificadas por plasmídeos, fagos, transposons e sequências de inserção [220], sugerindo que tenham sido adquiridas por transferência horizontal de genes [221].

O *spot* 7 (tab. 3) foi identificado como uma proteína hipotética conservada de função desconhecida, que contém resíduos de cisteínas agrupados. Algumas das proteínas que apresentam este domínio contêm sítios de ligação a Zn [222, 223] ou Cu e Fe-S [224]. Por sua vez, as enzimas que coordenam grupamentos Fe-S incluem aconitase e succinato desidrogenase (SDH) do TCA, a nitrogenase envolvida na fixação do nitrogênio e uma proteína que responde a Fe [224]. As primeiras evidências sobre a biogênese dos grupos Fe-S foram derivadas do estudo da montagem do metalogrupo Fe-S da nitrogenase de *Azotobacter vinelandii* [225].

As demais proteínas identificadas na fase E são: proteína da biossíntese de polissacarídeos (*spot* 8), fator de iniciação 5A (IF-5A, *spot*5) e peptidase de sinal do tipo I (SpaseI, *spot*16). Em todas as bactérias analisadas até hoje, SPaseI é essencial para a viabilidade celular [226]. No entanto, SPaseI aparentemente são degradadas por auto-clivagem, representando um obstáculo à sua caracterização [226].

4.2.6- Análise das proteínas induzidas na fase estacionária (S)

Dois grupos de proteínas especificamente induzidos na fase S do crescimento de *G. diazotrophicus* são as transportadoras e transferases (fig. 22, tab. 4). Como esperado para esta fase, as células não estão muito envolvidas em neossíntese, mas têm habilidade de responder ao estresse da limitação de nutrientes, pela redução das atividades anabólicas e bloqueio dos sistemas de geração de energia [227].

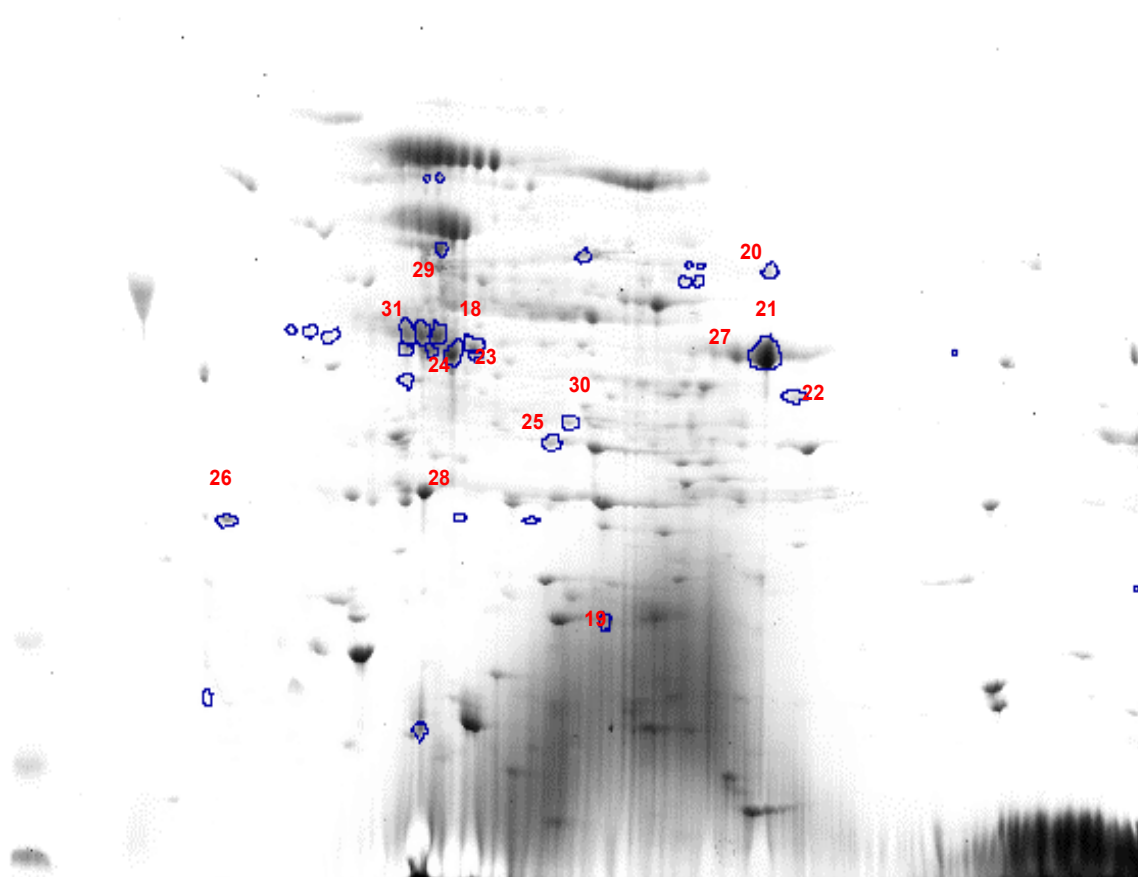


Figura 22: Perfil de 2D-GE de *G. diazotrophicus* em fase estacionária (S) de crescimento. Em azul estão destacados os *spots* induzidos nesta fase e os números em vermelho indicam as proteínas identificadas (tab. 4).

Uma dessas proteínas é a amidofosforibosiltransferase (*spot* 19, tab. 4 – produto do gene *purF*), a enzima limitante da via de síntese *de novo* de purinas [228].

Fosforibosiltransferases funcionais são essenciais para o crescimento celular, visto que o mutante *purF* de *M. smegmatis* 329B é incapaz de sobreviver em limitação de carbono [229]. Variações nos níveis de PurF ao longo da curva de crescimento, como observado em *G. diazotrophicus*, foram descritas na cianobactéria *Synechococcus* sp. cepa PCC 7942, cujo padrão de expressão de *purF* ocorreu em função do ritmo circadiano. Os autores sugeriram que este padrão de expressão diferencial pode estar relacionado com a sensibilidade ao oxigênio da amidofosforibosiltransferase [230]. Sugestivamente, detectamos PurF mais expressa em *G. diazotrophicus* na fase S, quando os níveis de O₂ dissolvido nas culturas é extremamente baixo (graf. 2).

Uma glicosiltransferase (*spot 22*) e uma proteína de transferência de multi-fosfatos (MTP, *spot 24*), ambas especificamente encontradas na fase S, estão envolvidas no metabolismo de açúcares. MTP foi primeiramente descrita em *R. capsulatus* como produto do gene *fruB* e envolvida no transporte de frutose para o citoplasma. MTP consiste de 3 domínios: 1) enzima III de especificidade a frutose, 2) enzima I (EI-like moiety) e 3) domínio FPr(HPr)-like moiety. Os domínios HPr e EI descritos na literatura apresentam homologia com membros do sistema PTS de transferência de fosfatos a açúcares [231]. Um desses homólogos, DhaM, foi encontrado em *E. coli*, porém não relacionado ao sistema PTS [232]. Recentemente, análise genômica de *Desulfovibrio vulgaris* identificou uma proteína MTP homóloga a DhaM com especificidade a manose [233]. A proteína MTP identificada em *G. diazotrophicus* precisa de caracterizações bioquímicas, mas, de acordo com resultados de estudos em outras bactérias, suas funções devem incluir fosforilação e captação de açúcares [231].

Foram encontrados 3 transportadores da família ABC (*ATP-Binding Cassette*) superexpressos na fase S. Um deles é o componente periplasmático do sistema de transporte de dipeptídeos (DppA, *spot 18*, tab. 4). Em *E. coli*, DppA capta dipeptídeos exógenos para o

citoplasma [234] e, curiosamente, os níveis do mRNA de *dppA* na célula dependem da fase de crescimento [235]. DppA já foi descrita em outras bactérias. Em *R. Sphaeroides*, além de sua função no transporte de dipeptídeos, DppA atua como uma chaperona molecular [236]. Em *Haemophilus influenzae* uma proteína ligadora de heme é homóloga a DppA [237] e, em bactérias entéricas e *Rhizobium leguminosarum* [238], DppA também funciona na importação de ácido δ -aminolevulínico (ALA), um precursor de heme com características estruturais a dipeptídeos [239, 240]. Como em *E. coli*, a DppA de *G. diazotrophicus* é expressa de forma dependente do crescimento e pode atuar no transporte de peptídeos e ALA, visto que o grupo heme é necessário para a alta demanda de energia da bactéria.

Outros membros da família ABC são subunidades ligadoras de ATP dos sistemas de transporte de ribose (*spot* 20, tab. 4) e de exportação de β -1,2 glucanas (*spot* 29, tab. 4). O sistema de transporte de alta afinidade de ribose em *E. coli* compreende uma componente periplasmática (RBP or RbsB), uma de membrana (RbsC) e uma proteína ligadora de ATP (RbsA). Este sistema transporta ribose em baixas concentrações de açúcares [241] e um conjunto similar de proteína foi identificado em *M. smegmatis*, especificamente na fase S [242]. A detecção de um homólogo RbsA (*spot* 20, tab. 4) em células da fase S sugerem a presença de um operon *rbsACB* homólogo ao de *E. coli* em *G. diazotrophicus*, possivelmente com função de transportar ribose para o interior da célula em condições de limitação de nutrientes.

A síntese e o transporte de glucanas cíclicas β -1,2 foram primeiramente descritas em *Agrobacterium* (ChvA), *Rhizobium* (NvdA) e *Brucella abortus* [243-245]. Mutantes destas espécies, deficientes na síntese e transporte dessas glucanas cíclicas apresentaram redução da infectividade e virulência [244, 246], sugerindo um papel fundamental dessas glucanas na interação dessas bactérias com seus hospedeiros. A detecção de uma proteína do sistema de

transporte de glucanas cíclicas em *G. diazotrophicus* é curiosa, embora sua função ainda seja desconhecida.

Bactérias respondem à deficiência de nutrientes no ambiente natural ou em cultura na fase S, um estado fisiológico caracterizado pela expressão de genes de resposta a estresse [113, 247, 248]. Nesta fase, *G. diazotrophicus* superexpressou uma leucil tRNA sintetase (*spot* 30, tab. 4), uma proteína envolvida na biogênese do leucil-aminoacil-tRNA, provavelmente como uma resposta a limitação do aminoácido leucina, conforme observado em outros microorganismos [249]. No entanto, aminoacil tRNA sintetases (aaRS) também são capazes de sintetizar diadenosina tetrafosfato (Ap₄A), uma alarmona associada à resposta a uma grande diversidade de estresses. [250, 251]. Análise quantitativa realizada por O'Farrel em géis de acrilamida, mostrou que 18 das 21 aminoacil-tRNA sintetases tem sua concentração aumentada de forma dependente da fase de crescimento celular [252].

A ribulose 1,5-bisfosfato carboxilase/oxigenase (RuBisCO, *spot* 31, tab. 4) também foi expressa diferencialmente na fase S. RuBisCO é a proteína de maior abundância na natureza e participa do ciclo de Calvin durante a fotossíntese, que converte CO₂ atmosférico em carbono orgânico [253]. Sua expressão por *G. diazotrophicus* é intrigante, embora não tão surpreendente. Diversas bactérias não fotossintéticas como espécies de *Bacillus*, e as diazotróficas *S. meliloti* e *M. loti* possuem genes homólogos ao da RuBisCO. Análise da sequência de aminoácidos desses genes permitiu sua classificação em 4 formas (I a IV) [254]. A forma IV é encontrada em *B. subtilis* e outras bactérias não fotossintéticas e parece ser a expressa por *G. diazotrophicus*. Contudo, essa forma não exibe atividade carboxilase nem mesmo oxigenase [255], mas na via de metionina [254], no metabolismo de enxofre e na resposta ao estresse oxidativo [256].

NADH:ubiquinona oxidoreductase (*spot* 25, tab. 4), outra proteína da fase S, por outro lado, é uma bomba de cátions acoplada ao complexo respiratório I, que transfere elétrons do

NADH para a ubiquinona com a translocação de cátions pela membrana de várias bactérias [257, 258]. Na fase S ainda foram encontradas, a Mg-protoporfirina IX quelatase, uma ATPase da superfamília com domínio pirofosfato PP-*loop* (*spot* 23, Fig. 3S), que apresenta domínios de ligação e hidrólise de nucleotídeos e pode estar envolvida em processos tais como biossíntese de cofatores e modificação de tRNA [259] e a proteína FixG (*spot* 21, tab. 4) cuja função especula-se na próxima seção.

4.2.7- Análise das proteínas induzidas em limitação de compostos nitrogenados (L)

A maioria das proteínas especificamente expressas em L aparentemente estão envolvidas com o processo de fixação de nitrogênio e a habilidade simbiótica de *G. diazotrophicus* (fig. 23, tab. 5). As proteínas do complexo da nitrogenase, NifH, NifD e NifK e NifA, esse último o ativador transcricional do operon *nifHDK*, não foram encontrados, possivelmente devido à sua baixa abundância na célula e à pouca sensibilidade do método de coloração por Coomassie.

Duas proteínas exclusivamente detectadas em L foram NodC (*spot* 39, tab. 5) e NodM (*spot* 52, tab. 5), que em *Rhizobium* são essenciais à formação do nódulo simbiótico [260-262]. NodC é uma N-acetil-glicosaminil transferase e NodM uma glicosamina-fructose-6-fosfato aminotransferase [260-262]. Ambas fazem parte do metabolismo de aminoaçúcares da parede celular e em *R. meliloti*, NodC está envolvida na especificidade do hospedeiro [263].

Outra proteína de envelope celular identificada em L é a proteína 2 ligadora de penicilina (PBP2, *spot* 35, tab. 5) [264, 265]. PBP é uma família de proteínas envolvidas nos estágios tardios da síntese de mureína [266], que mantém a integridade e forma das células [266]. A PBP2 de *E. coli* possui atividade transpeptidase e está envolvida no alongamento celular [267]. Sua mutação leva a formação de células esféricas [267]. De acordo com Muthukumarasamy e colaboradores [15] altos níveis de nitrogênio combinado no meio de

cultura, especialmente na forma de amônio, induzem alterações morfológicas em *G. diazotrophicus*, resultando em células longas e pleiomórficas. Essas evidências sugerem uma ligação entre PBP2 e, possivelmente NodC, NodM e uma lipoproteína associada a peptídeoglicanos (*spots* 45-49) que atuariam no remodelamento da parede celular de *G. diazotrophicus* durante o processo de fixação de nitrogênio.

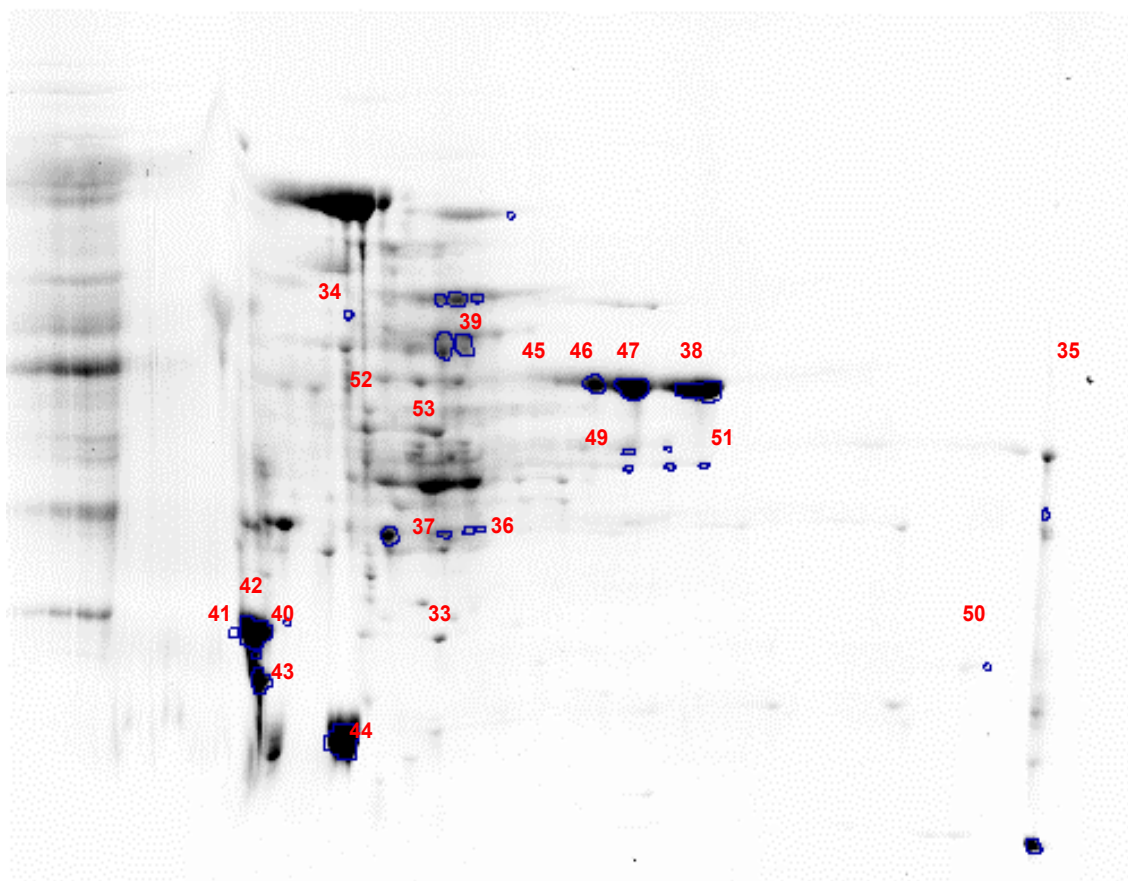


Figura 23: Perfil de 2D-GE de *G. diazotrophicus* em condição de limitação (L) de nitrogênio. Em azul estão destacados os *spots* induzidos nesta fase e os números em vermelho indicam as proteínas identificadas (tab. 5).

O regulador transcricional MatR, mais expresso em L, foi primeiramente descrito em *R. leguminosarum* *bv. trifolii* como um repressor do operon *matABC* operon, envolvido na captação e conversão de malonato a acetil-CoA [268, 269], [270]. Portanto, a expressão de MatR em *G. diazotrophicus* na condição L indica uma potencial estratégia da bactéria para utilização do malonato como fonte de carbono.

FixG é uma proteína que foi descrita em *Rhizobium sp* como produto do gene *fixG* do operon *fixGHI* [271, 272]. Aparentemente, esta proteína catalisa reações de oxidação-redução acopladas à FixI, a subunidade catalítica de uma bomba de prótons tipo P, requerida especificamente para fixação de nitrogênio simbiótica [271]. A expressão de FixG por *G. diazotrophicus* em uma condição de fixação de nitrogênio sugere um papel para esta proteína neste processo.

Outras proteínas típicas da fase L associadas ao processo de fixação de nitrogênio são: NifB (*spot* 50, tab. 5), NifW (*spot* 44, tab. 5) e NifR3 (*spot* 53, tab. 5) e HypA (*spot* 33, tab. 5) e HypX (*spot* 34, tab. 5). NifB, NifE e NifN estão envolvidas na biossíntese do cofator FeMo do complexo da nitrogenase [273]. NifW protege a FeMo-proteína da nitrogenase de danos oxidativos [274] e NifR3 é uma proteína conservada de função desconhecida. No entanto, foi relatado que *R. capsulatus* em limitação de nitrogênio aumenta a expressão de NifR3 mais de 5 vezes, sugerindo uma função no processo de fixação [275]. A expressão de diversas proteínas Nif em *G. diazotrophicus*, com funções conservadas entre várias bactérias, sugerem que elas funcionem de maneira semelhante em *G. diazotrophicus*. A expressão de duas hidrogenases, HypA e HypX, sugere que o processo de fixação de nitrogênio esteja acoplado à reciclagem de H₂ em *G. diazotrophicus*, como ocorre em outras espécies [276]. O hidrogênio liberado durante a fixação de nitrogênio pode ser reciclado por hidrogenases para aumentar a eficiência metabólica da diazotrofia [277].

A diene lactona hidrolase (DHL, *spot* 37, tab. 5), enzima de detoxificação responsável pela degradação de compostos aromáticos clorinados em *Pseudomonas*, *Rhodococcus opacus* e *Ralstonia eutropha* [278-280] também foi diferencialmente expressa em L pela *G. diazotrophicus*. DLH pertence a família das α/β -hidrolases [281] e, em *G. diazotrophicus*, pode atuar na detoxificação de subprodutos gerados no processo de fixação de nitrogênio e /

ou durante seu estágio simbiótico, lidando com os compostos aromáticos produzidos em abundância pelas plantas [282].

Outras proteínas expressas em L são uma proteína hipotética (*spots* 40-43) e um regulador anaeróbico (*spot* 32), cuja expressão em L pode ser uma resposta ao baixo nível de O₂ dissolvido no meio.

4.2.8- Análise das proteínas induzidas em abundância de compostos nitrogenados (H)

Apenas 4 proteínas induzidas na condição H foram identificadas (fig. 24, tab. 6). Duas são peptidases: aminopeptidase P (APP, *spot* 55, tab. 6) e peptidase de sinal I (SPaseI, *spot* 56, tab. 6). As outras são UreD (*spot* 57, tab. 6) e uma integrase (*spot* 54, tab. 6). APP e SPaseI são peptidases que especificamente removem aminoácidos N-terminais e atuam na exportação de proteínas. [226]. Por outro lado, integrases são proteínas associadas à recombinação do DNA [220].

UreD é uma proteína acessória envolvida na montagem de metalogrupos de Ni da apoproteína urease [283]. Ureases são complexos metaloenzimáticos que catalisam a hidrólise de uréia a amônia e dióxido de carbono, permitindo que muitas bactérias do solo utilizem-na como fonte de nitrogênio [284]. Sua expressão por *G. diazotrophicus* em H representa uma estratégia adicional para captação de compostos nitrogenados do ambiente e além disto, a produção de amônia pode funcionar como uma maneira de aumentar o pH ácido do meio, garantindo portanto uma proteção extra contra danos por acidez [195] como já observado em *Helicobacter pylori* [285].

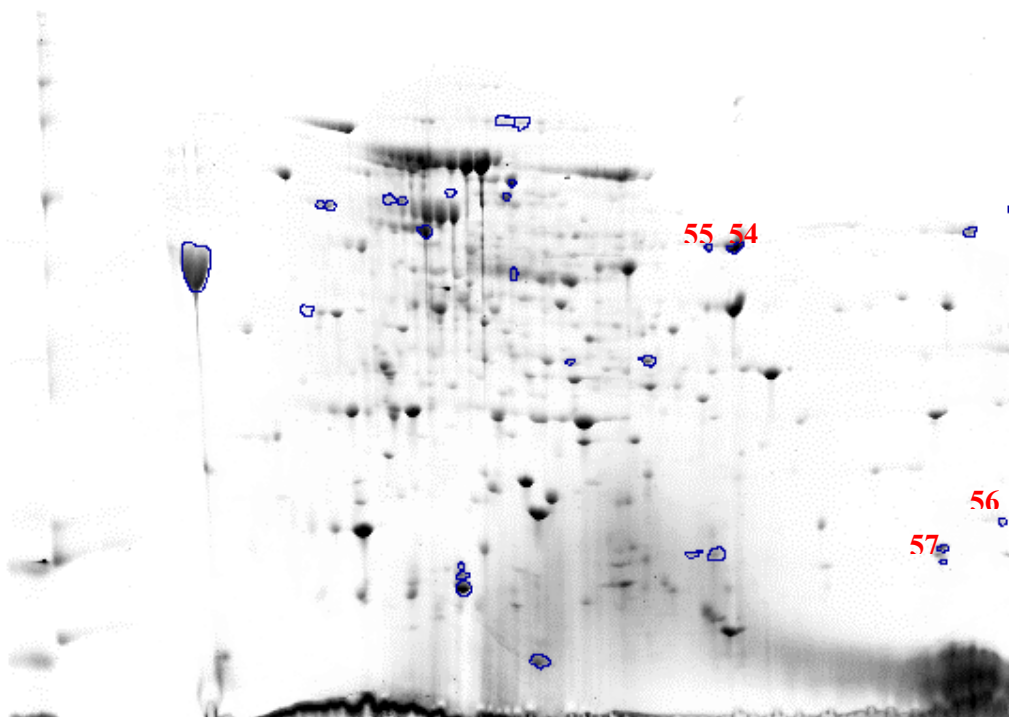


Figura 24: Perfil de 2D-GE de *G. diazotrophicus* em condição de abundância (H) de nitrogênio. Em azul estão destacados os *spots* induzidos nesta fase e os números em vermelho indicam as proteínas identificadas (tab. 6).

4.2.9- Conclusões parciais

- Foram obtidos mapas bidimensionais de *G. diazotrophicus* PAL5 em condições de limitação e abundância de compostos nitrogenados, nas fases exponencial e estacionária de crescimento.
- Foram identificadas 31 proteínas diferencialmente expressas entre as fases exponencial e estacionária do crescimento de *G. diazotrophicus* e 26 proteínas entre as condições de limitação e abundância de compostos nitrogenados. A identificação de proteínas relacionadas com o processo de fixação de nitrogênio, induzidas especificamente na condição de limitação de nitrogênio, sugere a ocorrência deste processo nas condições analisadas.
- A eletroforese bidimensional para obtenção de proteomas se confirmou como uma excelente ferramenta para observar processos celulares de uma forma global (possibilidade de

analisar quase 500 proteínas simultaneamente) e para identificar subconjuntos de proteínas (sub-proteomas) envolvidos em processos específicos;

- Foi estabelecida uma correlação direta entre grande parte das proteínas identificadas e as condições em que foram expressas;

- A análise de proteomas diferenciais apontou proteínas que merecem estudos moleculares e bioquímicos, para que seus potenciais biotecnológicos sejam avaliados;

- Experimentos futuros que utilizem análise de géis bidimensionais na faixa de pH 4 a 7 e a técnica DIGE, que utiliza a marcação fluorescente das proteínas antes da corrida, possibilitando a análise comparativa entre duas ou três amostras diferentes simultaneamente, poderão confirmar e acrescentar informações, considerando o poder de resolução desta tecnologia.

4.3- PROTEOMA X GENOMA

4.3- GENOMA X PROTEOMA

4.3.1 – Contribuições do proteoma para a anotação do genoma

Embora o número de proteínas identificadas por MudPit tenha sido muito maior do que de proteínas identificadas por gel 2D e MALDI, a análise por eletroforese bidimensional nos proporcionou informações que contribuíram substancialmente para a anotação das ORFs previstas pelo genoma.

A comparação entre valores de pI e MM obtidos para as proteínas resolvidas em gel 2D das proteínas identificadas com os valores teóricos, preditos a partir das ORFs deduzidas de suas sequências gênicas nos revelou uma discrepância significativa entre esses valores (fig. 25), em alguns casos diferenças maiores do que 10 Da e 2 unidades de pH foram observadas. Na figura 25 é possível observar que para muitas proteínas a MM prevista foi maior do que a observada em gel, assim como diversas proteínas identificadas tiveram valores de pI previstos maiores do que 7, que não são resolvidos nas tiras utilizadas.

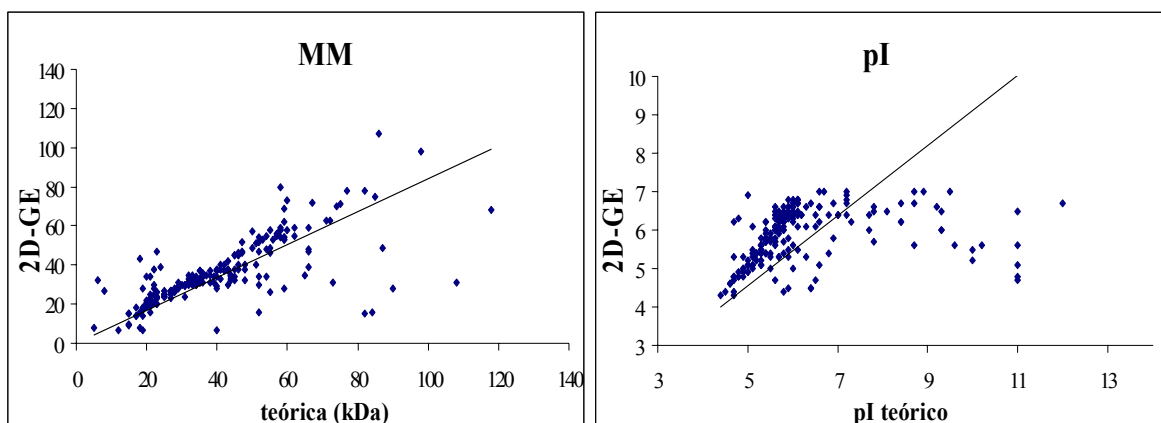


Figura 25: Comparação entre os valores experimentais e teóricos de pI e MM das proteínas identificadas pela estratégia gel 2D/MALDI.

Das 208 proteínas diferentes identificadas por 2D-MALDI, 96, ou seja, 45%, apresentaram pI e MM experimentais muito divergentes dos teóricos. Este resultado nos levou a considerar cinco principais hipóteses:

- 1- Identificação incorreta da proteína

- 2- Processamento pós-traducional da proteína
- 3- pI teóricos não corresponde ao pI da proteína
- 4- Clivagens das proteínas por proteases presentes no extrato
- 5- Previsão incorreta da ORF no genoma correspondente à proteína observada em gel

A primeira opção não nos pareceu muito provável pois um grande número de peptídeos foi utilizado na identificação de cada uma dessas proteínas e muitas dessas proteínas foram identificadas não apenas por *fingerprint* de massa, mas também por seqüenciamento de peptídeos com dados da fragmentação desses peptídeos.

Já os processamentos pós-traducionais mais comuns em bactérias, fosforilações e oxidações principalmente, não causam alteração significativa de MM, apenas pequenas variações de pI, o que também não justificaria grande parte das discrepâncias observadas. Ainda assim, verificamos que, das 45 proteínas com discrepância apenas entre valores de pI experimental e teórico, apenas 5 deveriam conter fosforilações, de acordo com o programa Scansite (<http://scansite.mit.edu/>). Outro tipo de processamento seria a clivagem de peptídeo sinal, predito pelo SignalP para 17 dessas proteínas, entretanto, 3 das 17 são sabidamente citoplasmáticas, ou seja, apenas 14 devem ter sido processadas desta forma, justificando assim as diferenças de pI observadas.

A hipótese de que algumas proteínas, cuja MM observada em 2D-GE foi menor do que a prevista, podem ter sido clivadas por proteases presentes no próprio extrato é real, visto que o único inibidor de proteases utilizado foi o PMSF, o qual bloqueia a ação apenas de serino e algumas cisteíno proteases [134]. A previsão teórica de pIs baseia-se apenas na composição de aminoácidos, sem levar em conta a exposição ou internalização de determinados resíduos, quando a proteína adquire um enovelamento tridimensional. Visto que no processo de focalização isoeletrica as proteínas da amostra não estão completamente

desnaturadas, o pI experimental pode diferir do pI teórico, explicando algumas diferenças observadas.

Por estas observações, decidimos voltar ao banco de dados utilizado para detecção das ORFs (Riogene) e buscar a sequência codificante de cada um dos peptídeos identificados para cada uma das 573 proteínas em relação ao banco. Esta análise foi realizada no programa Artemis [146], utilizando as sequências dos 8 *contigs* disponibilizados pelo projeto Riogene em outubro de 2006.

Nesta etapa, localizou-se no banco de dados a sequência codificante de cada peptídeo obtido pela análise de massas. Foram considerados a qualidade dos espectros de massa, a posição de cada peptídeo em relação à ORF prevista, o conteúdo de GC da ORF prevista, a presença de peptídeos não identificados no banco de ORFs nas vizinhanças da ORF prevista, os valores de pI e MM dos principais homólogos de cada ORF e os valores de pI e MM da proteína observada em gel 2D.

Uma digestão com tripsina gera peptídeos que necessariamente terminam em lisina ou arginina (não precedidas de prolina) e que são precedidos por uma lisina ou arginina. Numa digestão tríptica perfeita, os únicos peptídeos que não seguem essas duas regras são os peptídeos N e C terminais da proteína. O peptídeo N-terminal, por não ser precedido por outro, apenas segue a regra da terminação em lisina ou arginina. Por outro lado, o peptídeo C-terminal, o último da proteína, mesmo sendo precedido por lisina ou arginina, não necessariamente termina em lisina ou arginina. Dessa forma, a identificação de peptídeos semi-trípticos nos leva a determinação de início e final de proteína e conseqüentemente da determinação de *start* e *stop* codons, sendo esses dois dos maiores problemas dos programas de previsão de ORFS.

Na prática, em cada espectro de massas foram selecionados apenas os picos que apresentaram uma relação sinal / ruído maior do que 10 e resolução em torno de 7000. A

identificação desses picos, que correspondem a peptídeos, foi feita pelo programa Mascot utilizando as sequências dos 8 *contigs* seqüenciados e anotados pelo projeto Riogene. Quando utilizamos um banco de ORFs ou de sequências de proteína no Mascot, o resultado apresentado é a proteína identificada ou uma lista das proteínas mais prováveis. No caso do banco de DNA, nenhuma informação sobre qual proteína foi identificada é disponibilizada, pelo fato do banco de DNA não conter tais informações. Assim, nas análises descritas a seguir, utilizamos o banco das sequências de DNA (dos 8 *contigs*) para identificação no Mascot, entretanto os dados do banco de ORFs foram acoplados utilizando o Artemis.

O Artemis é uma ferramenta de visualização de sequências de DNA e de resultados de programas de análise de sequências, tais como o FASTA e o BLAST, permitindo uma visão integrada das informações disponíveis. Como entrada para o Artemis foram utilizados: a) um arquivo no formato de texto que contém a sequência dos 8 *contigs*, b) um arquivo no formato EMBL, que contém as informações de localização no *contig*, sequência e anotação de cada ORF (resultados do BLAST e PFAM, por exemplo) e c) um arquivo no formato EMBL gerado pelo Mascot contendo informações sobre cada peptídeo identificado, tais como o escore do peptídeo, sua sequência codificante e localização. Todas essas informações foram visualizadas simultaneamente, permitindo-nos conferir a anotação de cada ORF e a posição de cada peptídeo em relação às ORFs previstas. Adicionalmente, o Artemis apresenta um gráfico do conteúdo de GC ao longo da sequência de DNA. Este gráfico é uma fonte importante de informação, visto que em geral, em bactérias, as regiões codificantes apresentam conteúdo de GC mais alto do que as regiões não-codificantes. Quando necessário, para visualização dos alinhamentos das sequências, os programas BLAST e COG foram utilizados diretamente da internet, através dos sítios <http://www.ncbi.nlm.nih.gov/blast> e <http://www.ncbi.nlm.nih.gov/cog>.

Estas análises permitiram a validação experimental das 573 ORFs que codificam as proteínas identificadas e algumas correções a saber (tab. 7):

- a. Corrigir o início da transcrição (*start codon*) de 31 N-terminais pelo seqüenciamento de peptídeos semi-trípticos;
- b. Corrigir o fim da transcrição (*stop codon*) de 25 C-terminais pelo seqüenciamento de peptídeos semi-trípticos;
- c. Alterar o N ou C –terminal de 16 proteínas previstas com MM menores do que deveriam;
- d. Alterar o N ou C –terminal de 12 proteínas previstas com MM maiores do que deveriam;
- e. Alterar o *frameshift* de 5 ORFs previstas erroneamente;
- f. Validar 3 ORFs melhor previstas numa versão anterior do banco do genoma;
- g. Assinalar 13 ORFs não previstas pelo projeto genoma.

Tabela 7: Contribuições do proteoma para a anotação do genoma de *G. diazotrophicus*,

Confirmação experimental de ORF	573
Determinação experimental do N-terminal por sequenciamento	31
Determinação experimental do C-terminal por sequenciamento	25
Sugestão de aumento no tamanho da ORF (MW gel e homólogos)	16
Sugestão de diminuição no tamanho da ORF (MW gel e homólogos)	12
Deteccção de erros de <i>frameshift</i> no sequenciamento	5
Substituições pontuais	3
Assinalamento de nova ORF	13

Na tabela 1 do anexo pode-se verificar as contribuições dessas análises proteômicas para a anotação de cada uma das proteínas identificadas. Para ilustrar essas análises, tem-se o caso de uma proteína (*spot c104*) cuja ORF foi prevista menor do que a ORF real (figs. 12 e 13). A análise dos dados de MS/MS do *spot c104* no banco de ORFs da *G. diazotrophicus*, através do programa Mascot, não resultou em identificação da proteína (fig. 26A). Na página do resultado da identificação, o Mascot apresenta um gráfico do número de *hits* por escore. Um bom resultado é indicado pela presença de uma única barra vermelha, fora da área hachurada em verde, indicando uma proteína com probabilidade de identificação maior do que 95%. Uma nova busca pela identificação das sequências codificantes dos peptídeos desta proteína no banco de sequências de DNA da *G. diazotrophicus* resultou numa identificação possível (fig. 26B). Ou seja, quando o banco de sequências de DNA foi utilizado, ao invés do banco de ORFs, a identificação foi possível em um dos *contigs*. A visualização destes

resultados no Artemis mostrou a identificação de 12 peptídeos da proteína do *spot* c104 numa mesma região do contig. Ainda no Artemis, observou-se que esta região é parcialmente ocupada por uma ORF. Na verdade, 5 picos do espectro de massas correspondem à peptídeos da ORF e outros 7 picos correspondem à peptídeos da região adjacente à tal ORF. Desses 7 peptídeos, 4 foram seqüenciados. Toda a região que engloba os 12 peptídeos (7 peptídeos adjacentes + 5 peptídeos dentro da ORF) possui alto conteúdo de GC, sugerindo que seja uma região codificante (fig. 27A). A ORF em questão apresenta MM de 18 kDa e pI 4,8, enquanto que seus homólogos (análise dos resultados do BLASTp) apresentam massa 43 kDa e pI 6,3 e os valores experimentais do *spot* que gerou esta identificação são 41 kDa e pI 5,7 (fig. 27). Todas essas evidências nos levaram a sugerir a extensão da ORF. Na nova proposta, a ORF ficaria com MM 42 kDa e pI 6,0, totalmente dentro da região de alto conteúdo de GC e compatível com todos os dados experimentais e de bioinformática obtidos.

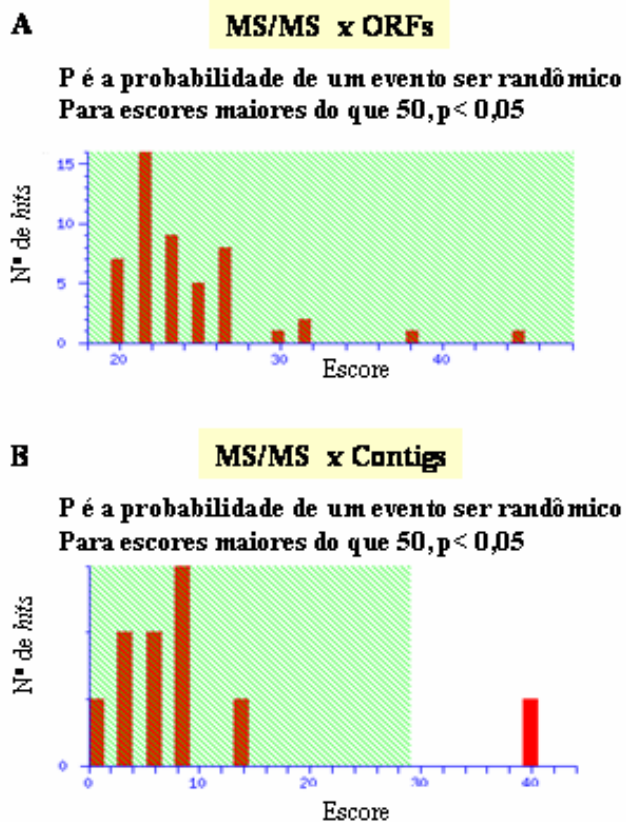


Figura 26: Exemplo de uma proteína (*spot c104*) cuja identificação proteômica contribuiu para anotação da ORF correspondente (ORF foi prevista menor do que a ORF real). Quando os dados de MSMS desta proteína foram comparados ao banco de ORFs, não resultou em identificação da proteína (A), mas quando comparados ao banco de DNA a identificação foi possível com $p < 0,05$, o que pode ser observado pela presença de uma barra vertical vermelha, fora da região hachurada em verde (B).

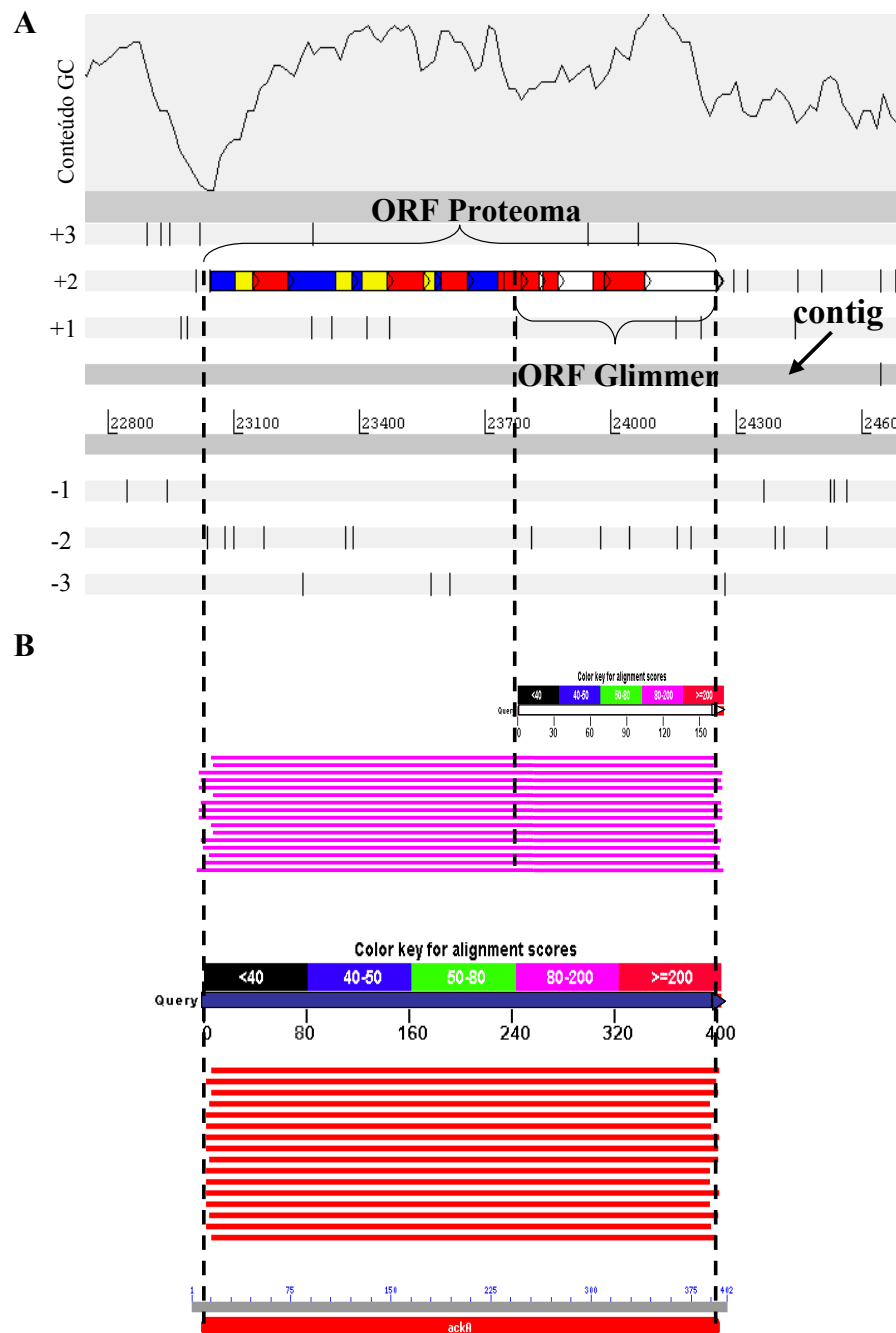


Figura 27: A ORF prevista pelo Glimmer, correspondente a proteína do *spot c104*, encontra-se dentro de uma região de alto conteúdo de GC (A). Em vermelho estão marcados os peptídeos do espectro de massas do *spot c104* identificados por *fingerprint* e em amarelo os peptídeos seqüenciados por MS/MS (A). Essa ORF apresenta cerca de 190 aminoácidos (MM de 18 kDa e pI 4.8), enquanto que seus homólogos apresentam aproximadamente 400 aminoácidos (MM em torno de 43 kDa e pI 6.3) (B).

Um outro exemplo é o da ORF 4_00623, que em contraste com o que ocorreu com o *spot* c104, foi inicialmente prevista maior do que a ORF real. Esta ORF maior apresenta-se no *frame* -1 de leitura e, de acordo com a previsão de ORFs do Glimmer, sobrepunha-se inteiramente à ORF 4_00624, codificada no *frame* +1 (fig. 28A). O espectro de massas do *spot* c185 identificou 9 peptídeos diferentes, todos na região C-terminal da ORF 4_00623, sendo um deles um peptídeo semi-tríptico, sugerindo que o N-terminal da proteína corresponda a este peptídeo (fig. 28). Adicionalmente, as sequências protéicas homólogas e ortólogas às ORFs 4_00623 e 4_00624 são substancialmente menores do que as previstas para estas ORFs. A ORF 4_00623 possui homólogos e ortólogos apenas no seu C-terminal, justamente a região que se localizam os peptídeos do espectro de massas, enquanto que a ORF 4_00624 também apresenta homólogos e ortólogos apenas no seu C-terminal (fig. 28). A ORF 4_00623 inicial apresenta MM maior do que 100 kDa e pI 7,3. Por outro lado, seu homólogos possuem MM em torno de 30 kDa e pI aproximadamente 5 e o *spot* c185 MM de 30 e pI em torno de 6 (fig. 28). Com essas evidências sugerimos a alteração da ORF 4_00623 para uma ORF de MM 32 e pI 5.8, iniciada no peptídeo semi-tríptico identificado (fig. 28).

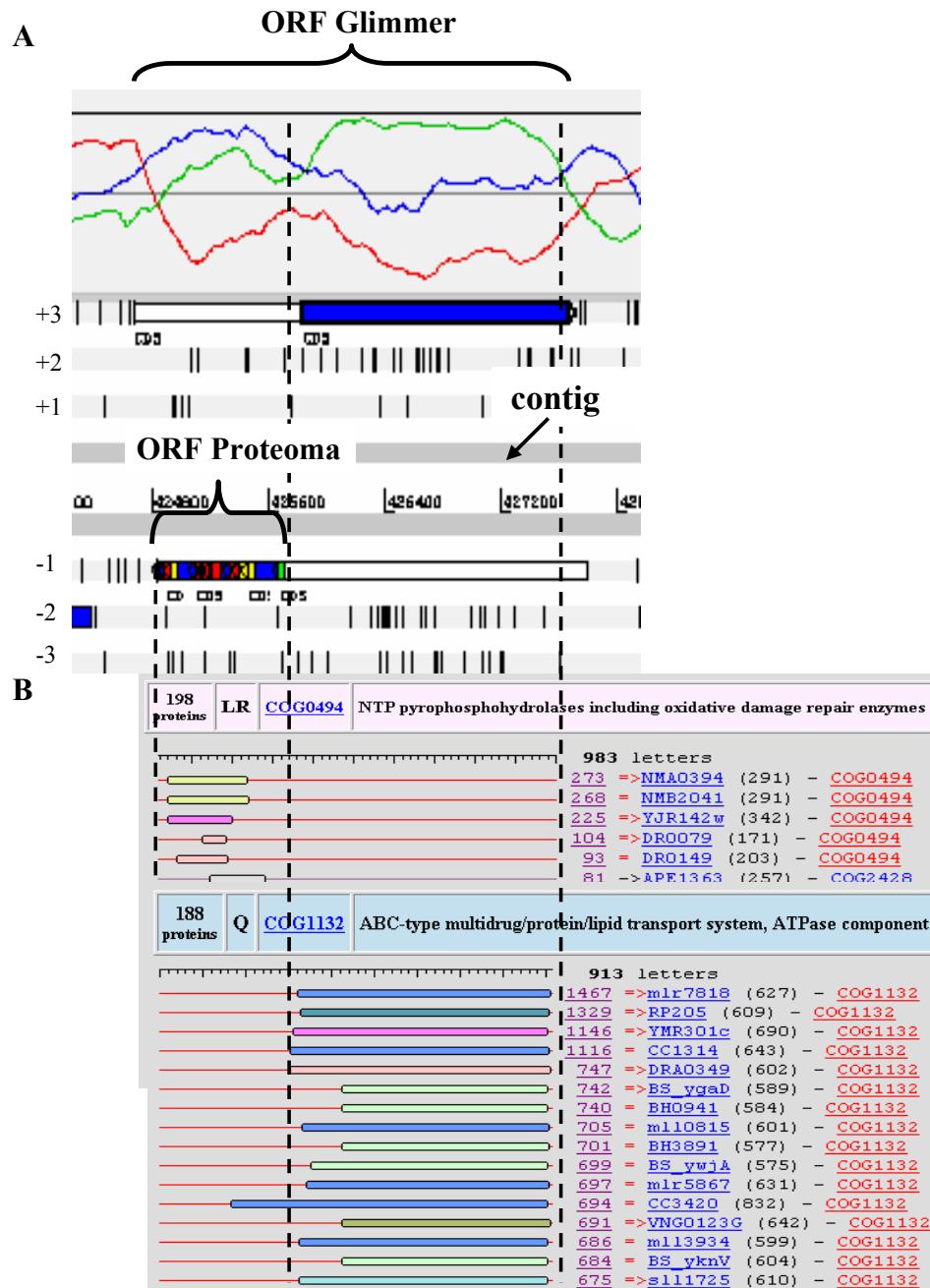


Figura 28: A) Região genômica da ORF 4_00623, mostrando ORFs previstas pelo Glimmer (branco + azul), peptídeos identificados por *fingerprint* (vermelho), peptídeos seqüenciados por MS/MS (amarelo), peptídeo semi-trípico N-terminal (verde) e ORFs propostas pela análise proteômica (azul). B) resultados do COG pra as ORFs 4_00623 e 4_00624.

Outra contribuição significativa desta análise proteômica para anotação do genoma foi a detecção de erros no seqüenciamento que causam *frameshifts*. Isso acontece, por exemplo, quando uma ou duas bases são perdidas na leitura da sequência, alterando o *frame* de leitura codificante. Quando isto ocorre, o programa que prediz as ORFs, no caso o Glimmer, termina a ORF prematuramente e inicia uma nova ORF no próximo *frame*. Com isso, a ORF fica dividida e as duas ORFs ficam incorretas: uma com seu C-terminal errado e a outra com o C e o N-terminal errados. No caso do *spot c3*, o resultado do Mascot identificou 2 proteínas com $p < 0,05$, produtos de ORFs consecutivas, uma no *frame* +1 e outra no +2. Ambas identificações pareciam confiáveis, apresentando altos valores de cobertura e número de peptídeos do espectro de massa identificados. A possibilidade de serem duas ORFs codificantes de proteínas similares, que apresentam domínios funcionais parecidos foi descartada. Tendo em vista a inexistência de *spots* nas adjacências do *spot c3* (o que poderia ser um indicativo da presença de duas proteínas no mesmo *spot*) e que os produtos das 2 ORFs eram homólogos a uma mesma proteína (BLASTp), sugerimos que essas ORFs sejam uma unidade única codificante, correspondendo as características da proteína correspondente ao *spot c3* (fig. 29).

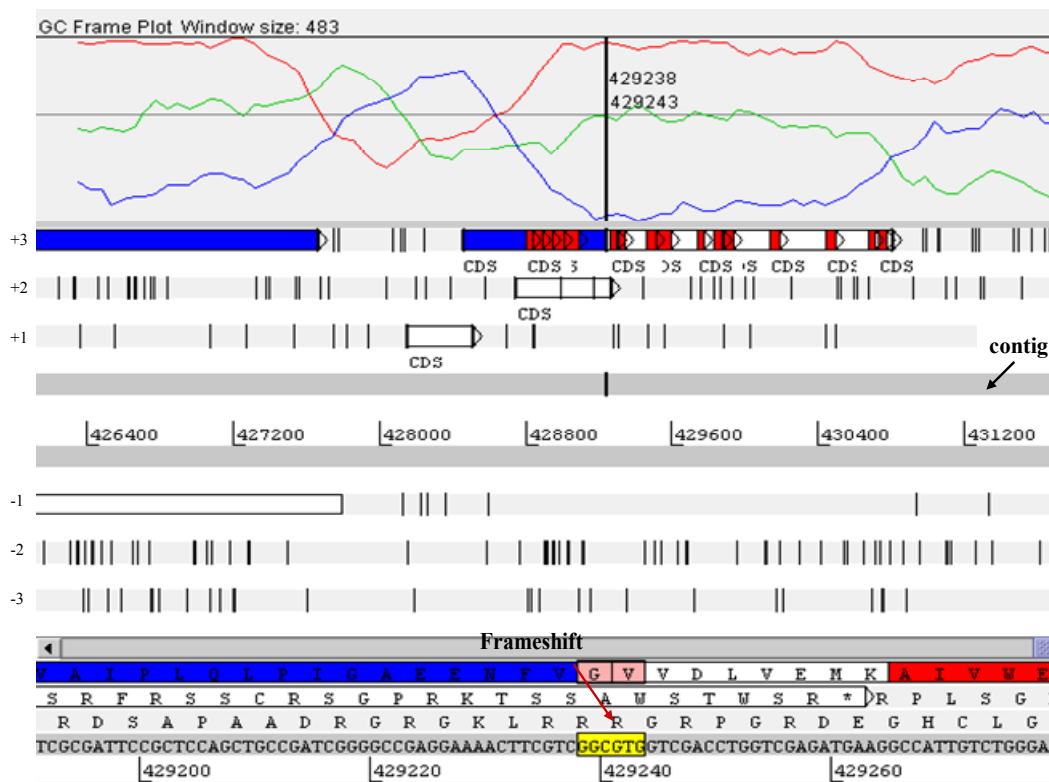


Figura 29: Exemplo da correção de um erro de *frameshift*. Região genômica da ORF 4_00627, mostrando ORFs previstas pelo Glimmer (branco), peptídeos identificados por *fingerprint* (vermelho), região do *frameshift* (amarelo) e ORFs propostas pela análise proteômica (azul).

Em diversos casos, foram encontrados peptídeos semi-trípticos correspondentes ao N ou C-terminal dos produtos de ORFs anotadas pelo consórcio Riogene. Em muitos deles, tais identificações apenas confirmaram o início e final das ORFs previstas, mas em outros casos esses peptídeos não coincidiram com os das ORFs previstas, permitindo-nos, portanto, sugerir pequenas alterações nas ORFs. Um desses casos está ilustrado na figura 30.

Por fim, quando cruzamos os espectros de massas obtidos com a sequência de DNA de *G. diazotrophicus*, algumas sequências de peptídeos localizaram-se fora das regiões previstas codificantes. Como sabemos que os genomas bacterianos, ao contrário dos eucarióticos, são tipicamente “econômicos”, não apresentando grandes regiões não-codificantes, pudemos sugerir a criação de novas ORFs. Cada caso foi analisado de forma independente e alguns

peptídeos fora de ORFs foram considerados apenas coincidência, sendo parte dos 5% de falsos positivos apontado pelo Mascot. Dentre as novas ORFs sugeridas (figs. 31 e 32), uma delas é especial, pelo fato das sequências codificadoras dos peptídeos em questão estarem sobrepostas a uma ORF hipotética. Neste caso (fig. 31), dois peptídeos seqüenciados por Q-TOF, cujas sequências genômicas correspondentes somente apresentam-se nessa região do genoma de *G. diazotrophicus*, foram localizados no *frame* +1. Nesta mesma região, no *frame* -2, foi prevista uma ORF. No entanto, nem a sequência do *frame* +1, nem do -2, possuem homólogos descritos até o momento. Neste caso, como os peptídeos não foram identificados através de uma proteína resolvida por gel 2D, não temos informações sobre MM e pI que pudessem ajudar a confirmar a identificação. Contudo, a previsão inicial de ORFs pelo Glimmer também apontava a possibilidade de uma ORF no *frame* +1, entretanto, esta foi removida durante o processo de anotação manual do genoma, tendo em vista a baixa probabilidade da ocorrência de duas ORFs sobrepostas em genomas de bactérias. Assim, sugerimos a deleção da ORF hipotética do *frame* -2 e a criação de uma nova ORF no *frame* +1.

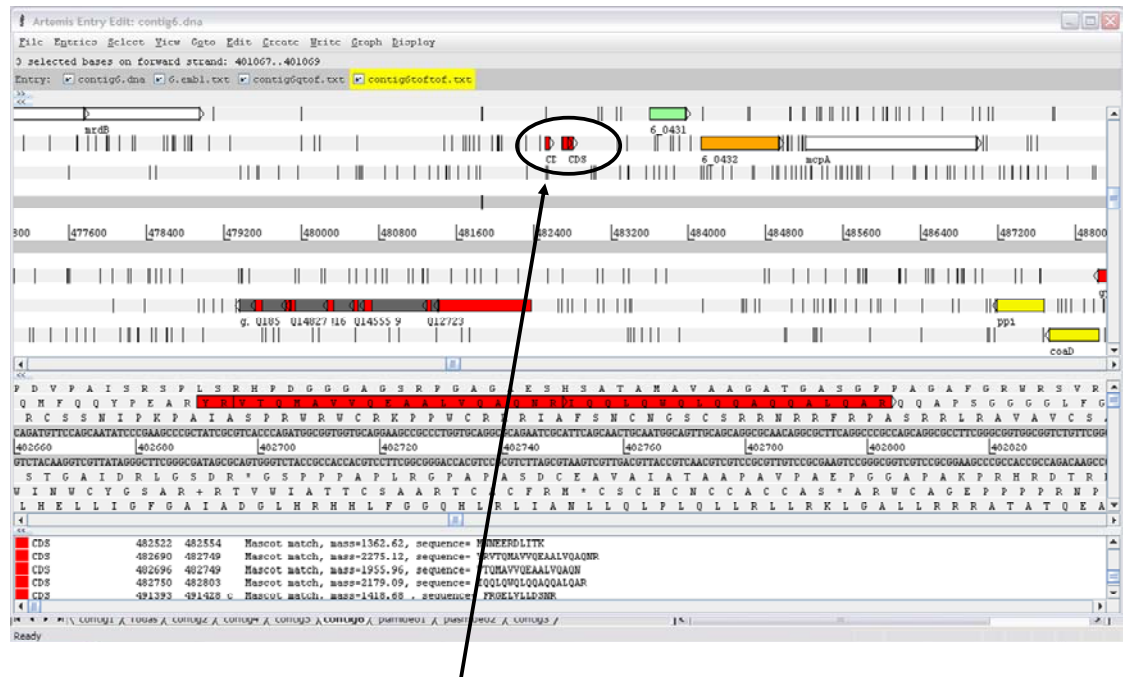


Figura 32: Região de 1.2 kb sem nenhuma CDS e quatro peptídeos de MS. Proposta de nova Orf, não prevista pelo Glimmer.

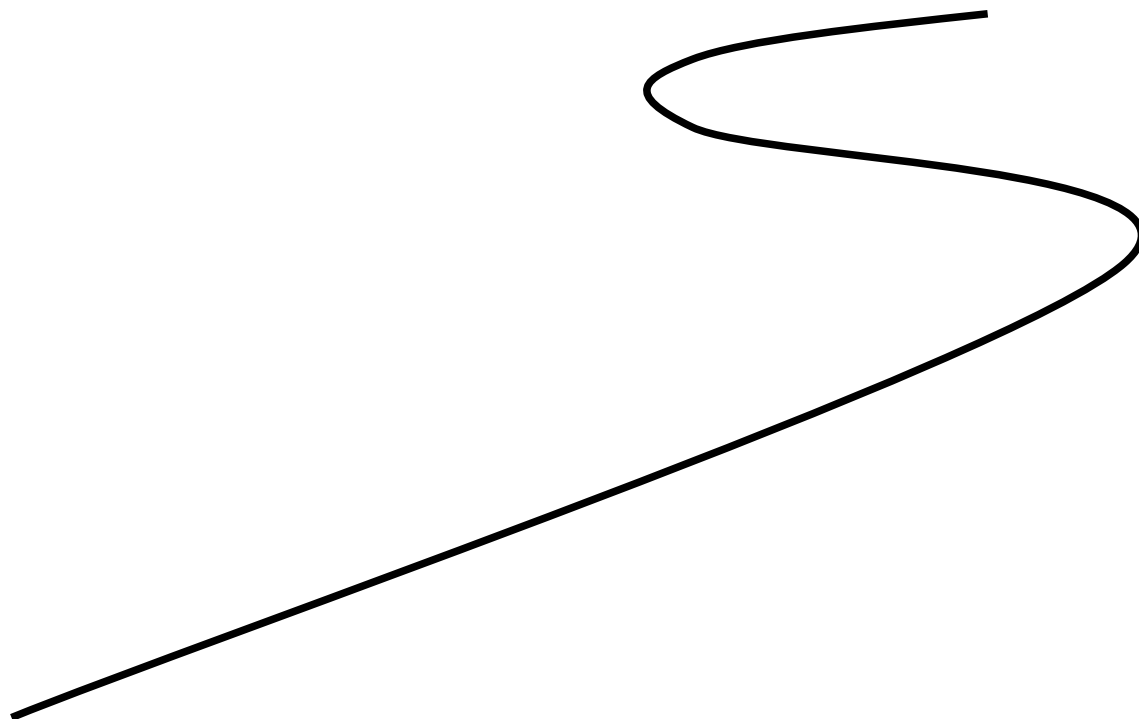
A utilização de dados de uma análise proteômica para o aprimoramento da anotação de genomas tem sido utilizada por diversos grupos. Embora recente, esta aplicação da espectrometria de massas é promissora, tendo revelado um grande número de novas ORFs no genoma de bactéria [286] e eucariotos [98, 287]. Nestes organismos, o cruzamento de espectros de massa com sequências brutas de DNA tem permitido a determinação de padrões de *splicing* alternativo, a definição de éxons e íntrons [288, 289].

4.3.2 – Conclusões parciais

Os dados de espectrometria de massas obtidos via 2D-GE ou MudPit para a construção do banco proteômico foram cruzados com o banco de ORFs e com o banco de sequências de DNA do projeto genoma da *G. diazotrophicus* (Riogene). Embora menos usual, a identificação das proteínas usando o banco de sequências de DNA trouxe informações

valiosas, permitindo a confirmação experimental e/ou correção de deficiências dos métodos puramente preditivos.

5- CONCLUSÕES



5- CONCLUSÕES

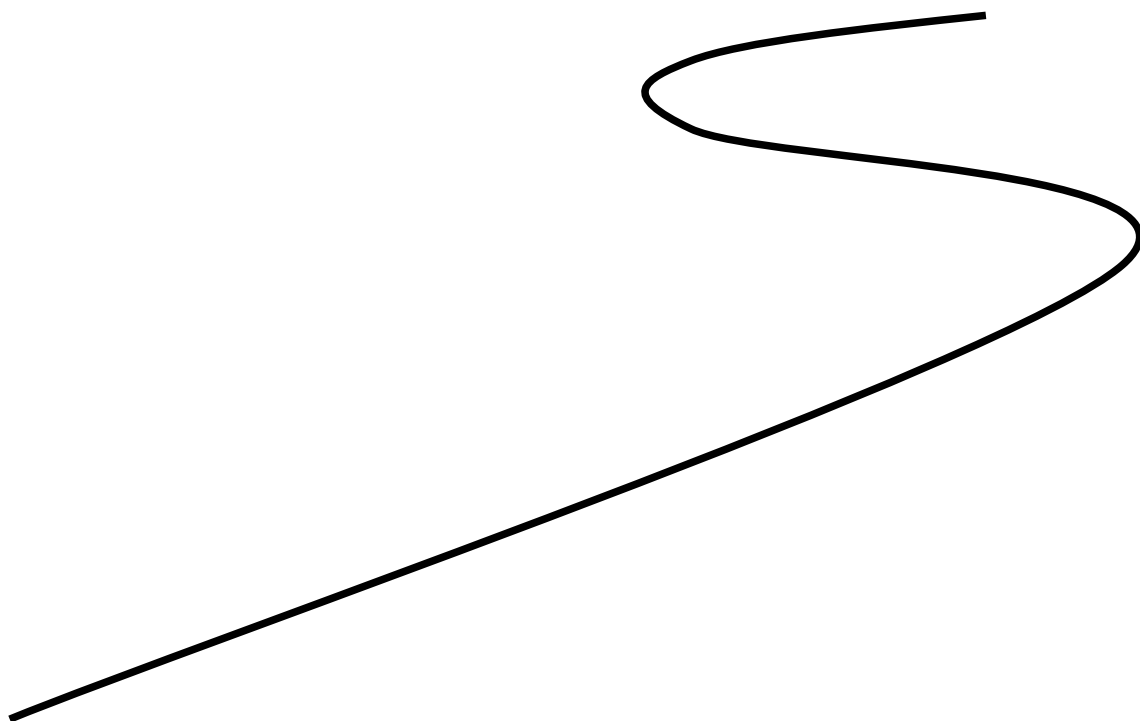
Este trabalho apresentou a primeira identificação em larga escala das proteínas expressas por *G. diazotrophicus*, tendo sido identificadas 208 proteínas a partir de eletroforese bidimensional e 526 pela estratégia de MudPit. Ao todo, 573 proteínas diferentes foram descritas e 40 vias metabólicas ativas foram detectadas. Essas vias metabólicas encontradas representam o metabolismo básico de uma célula bacteriana, compreendendo as principais vias do metabolismo de nucleotídeos, carboidratos, lipídeos, aminoácidos, cofatores e vitaminas, metabólitos secundários e de energia. A detecção da via do metabolismo de glioxilato e dicarboxilato é um outro fato novo e de destaque, visto que trabalhos anteriores com *G. diazotrophicus* não detectaram a atividade desta via [20]. Nenhuma via da biossíntese e metabolismo de glicanas e da degradação de xenobióticos foi detectada.

Ainda, 57 proteínas expressas diferencialmente e associadas especificamente com uma fase de crescimento, E ou S, ou com a disponibilidade de nitrogênio, L ou H, foram identificados, revelando diversas funções metabólicas tais como a homeostasia do pH citoplasmático, respostas a estresses, vias de degradação de compostos aromáticos, entre outras.

Por fim, em relação às análises por espectrometria de massas, mostramos que esta técnica, principalmente associada ao seqüenciamento de novo de peptideos, é uma ferramenta fundamental para a identificação de proteínas de organismos com genoma ainda não finalizado. A espectrometria de massas quando associada à 2D-GE deve ser vista não apenas como uma ferramenta para identificação de proteínas, mas principalmente para a caracterização, revelando modificações pós-traducionais, importantes na regulação metabólica de uma célula. Esta tecnologia pode ser ainda um instrumento adicional para a anotação de genomas, pela procura, nos espectros, por peptideos semi-trípticos para a detecção exata do N e C-terminais de proteínas e assim a delimitação precisa de ORFs.

Resumindo, neste trabalho foi dado o primeiro passo para a construção de um banco das proteínas expressas por *G. diazotrophicus* PAL5. Com a 2D-GE fizemos a identificação de 208 *spots*, dando início também a construção de um mapa de referência. Com a estratégia de MudPit foi possível identificar 526 proteínas de *G. diazotrophicus*. Portanto, este banco proteômico aqui apresentado contém 573 proteínas diferentes, o que não parece muito, mas na realidade é significativo, considerando que o genoma da bactéria não está completamente fechado e que algumas restrições metodológicas foram encontradas. A ampliação deste banco será possível pela análise de amostras de *G. diazotrophicus* cultivada em diferentes condições, possibilitando a expressão de genes específicos de uma condição de crescimento, a análise dos peptídeos do lisado celular, análise do sobrenadante de cultura e pela utilização de novas estratégias de fracionamento das proteínas, como por exemplo, a utilização de colunas de cromatografia por troca iônica. Ainda, pretendemos montar um sistema integrado e interativo de visualização e análise do banco de proteínas de *G. diazotrophicus*, o qual possa ser acessado via internet, atualizado regularmente e disponibilize todas as informações experimentais sobre tais proteínas .

REFERÊNCIAS BIBLIOGRÁFICAS



REFERÊNCIAS BIBLIOGRÁFICAS

1. Cavalcante, V.A. e Döbereiner, J., *A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane*. Plant and Soil, **108** (1): p. 23-31. 1988.
2. Yamada, Y., Hoshino, K. e Ishikawa, T., *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** (8): p. 1244-51. 1997.
3. Fuentez-Ramirez, L.E., Jimenez-Salgado, T., Abarca-Ocampo, I.R., *et al.*, *Acetobacter diazotrophicus*, an indoleacetic acid producing bacterium isolated from sugarcane cultivars of México. Plant and Soil, **154** (2): p. 145-150. 1993.
4. Li, R. e Macrae, I., *Specific association of diazotrophic Acetobacters with sugarcane*. Soil Biology and Biochemistry, **23**: p. 999-1002. 1991.
5. Ashbolt, N.J. e Inkerman, P.A., *Acetic Acid Bacterial Biota of the Pink Sugar Cane Mealybug, Saccharococcus sacchari, and Its Environs*. Appl Environ Microbiol, **56** (3): p. 707-712. 1990.
6. Döbereiner, J., Reis, V.M., Paula, M.A., *et al.*, *Endophytes diazotrophs in sugar cane, cereals and tuber plants*. In R. Palacios, J. Mora and W. E. Newton (ed.), *New horizons in nitrogen fixation*. Kluwer Academic Publishers. Dordrecht: p. 671-676. 1993.
7. Reis, V.M., Olivares, F.L. e Döbereiner, J., *Improved methodology for isolation of Acetobacter diazotrophicus and confirmation of its endophytic habitat*. World J. Microbiol. Biotechnol., **10**: p. 401-405. 1994.
8. Tapia-Hernandez, A., Bustillos-Cristales, M.R., Jimenez-Salgado, T., *et al.*, *Natural Endophytic Occurrence of Acetobacter diazotrophicus in Pineapple Plants*. Microb Ecol, **39** (1): p. 49-55. 2000.
9. Jimenez-Salgado, T., Fuentes-Ramirez, L.E., Tapia-Hernandez, A., *et al.*, *Coffea arabica L., a new host plant for Acetobacter diazotrophicus, and isolation of other nitrogen-fixing acetobacteria*. Appl Environ Microbiol, **63** (9): p. 3676-83. 1997.
10. Muthukumarasamy, R., Cleenwerck, I., Revathi, G., *et al.*, *Natural association of Gluconacetobacter diazotrophicus and diazotrophic Acetobacter peroxydans with wetland rice*. Syst Appl Microbiol, **28** (3): p. 277-86. 2005.
11. Loganathan, P., Sunlta, R., Parlda, A.K., *et al.*, *Isolation and characterization of two genetically distant groups of Acetobacter diazotrophicus from a new host plant Eleusine coracana L.* J. Appl. Microbiol., **87**: p. 167-172. 1999.
12. Youssef, H.H., Fayez, M., Monib, M., *et al.*, *Gluconacetobacter diazotrophicus: a natural endophytic diazotroph of Nile Delta sugarcane capable of establishing an endophytic association with wheat*. Biol Fert Soils, **39**: p. 391-397. 2004.
13. Paula, M.A., Reis, V.M. e Döbereiner, J., *Interactions of Glomus clarum with Acetobacter diazotrophicus in infection of sweet potato, sugar cane and sweet sorghum*. Biology and Fertility of Soils, **11**: p. 111-115. 1991.
14. Muthukumarasamy, R., Govindarajan, M., Vadivelu, M., *et al.*, *N-fertilizer saving by the inoculation of Gluconacetobacter diazotrophicus and Herbaspirillum sp. in micropropagated sugarcane plants*. Microbiol Res, **161** (3): p. 238-45. 2006.
15. Muthukumarasamy, R., Revathi, G., Seshadri, S., *et al.*, *Gluconacetobacter diazotrophicus (syn. Acetobacter diazotrophicus), a promising diazotrophic endophyte in tropics*. Current Science, **83** (2): p. 137-145. 2002.
16. Franke, I.H., Fegan, M., Hayward, A.C., *et al.*, *Nucleotide sequence of the nifH gene coding for nitrogen reductase in the acetic acid bacterium Acetobacter diazotrophicus*. Lett Appl Microbiol, **26** (1): p. 12-6. 1998.

17. Dong, Z., Canny, M.J., McCully, M.E., *et al.*, *A Nitrogen-Fixing Endophyte of Sugarcane Stems (A New Role for the Apoplast)*. *Plant Physiol*, **105** (4): p. 1139-1147. 1994.
18. Dobereiner, J., Baldani, V. e Baldani, J., *Como caracterizar e identificar bactérias diazotróficas de plantas não-leguminosas*. 1995, EMBRAPA Agrobiologia: Seropédica. p. 60.
19. Fuentes-Ramirez, L.E., Bustillos-Cristales, R., Tapia-Hernandez, A., *et al.*, *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** (Pt 4): p. 1305-14. 2001.
20. Alvarez, B. e Martínez-Drets, G., *Metabolic characterization of *Acetobacter diazotrophicus**. *Can. J. Microbiol.*, **41**: p. 918-924. 1995.
21. Stephan, M.P., Oliveira, M., Teixeira, K.R.S., *et al.*, *Physiology and dinitrogen fixation of *Acetobacter diazotrophicus**. *FEMS Microbiology Letters*, **77** (1): p. 67-72. 1991.
22. Luna, M.F., Bernardelli, C.E., Galar, M.L., *et al.*, *Glucose metabolism in batch and continuous cultures of *Gluconacetobacter diazotrophicus* PAL 3*. *Curr Microbiol*, **52** (3): p. 163-8. 2006.
23. Hernandez, L., Arrieta, J., Menendez, C., *et al.*, *Isolation and enzymic properties of levansucrase secreted by *Acetobacter diazotrophicus* SRT4, a bacterium associated with sugar cane*. *Biochem J*, **309** (Pt 1): p. 113-8. 1995.
24. Arrieta, J., Hernandez, L., Coego, A., *et al.*, *Molecular characterization of the levansucrase gene from the endophytic sugarcane bacterium *Acetobacter diazotrophicus* SRT4*. *Microbiology*, **142** (Pt 5): p. 1077-85. 1996.
25. Batista, F.R., Hernandez, L., Fernandez, J.R., *et al.*, *Substitution of Asp-309 by Asn in the Arg-Asp-Pro (RDP) motif of *Acetobacter diazotrophicus* levansucrase affects sucrose hydrolysis, but not enzyme specificity*. *Biochem J*, **337** (Pt 3): p. 503-6. 1999.
26. Hernandez, L., Arrieta, J., Betancourt, L., *et al.*, *Levansucrase from *Acetobacter diazotrophicus* SRT4 is secreted via periplasm by a signal-peptide-dependent pathway*. *Curr Microbiol*, **39** (3): p. 146-52. 1999.
27. Trujillo, L.E., Arrieta, J.G., Dafhnis, F., *et al.*, *Fructo-oligosaccharides production by the *Gluconacetobacter diazotrophicus* levansucrase expressed in the methylotrophic yeast *Pichia pastoris**. *Enzyme Microb Technol*, **28** (2-3): p. 139-144. 2001.
28. Martinez-Fleites, C., Ortiz-Lombardia, M., Pons, T., *et al.*, *Crystal structure of levansucrase from the Gram-negative bacterium *Gluconacetobacter diazotrophicus**. *Biochem J*, **390** (Pt 1): p. 19-27. 2005.
29. Reinhold-Hurek, B. e Hurek, T., *Life in grasses: diazotrophic endophytes*. *Trends Microbiol*, **6** (4): p. 139-44. 1998.
30. James, E.K., *Nitrogen fixation in endophytic and associative symbiosis*. *Field Crops Research*, **65**: p. 197-209. 2000.
31. Bellone, C., Bellone, D., Pedraza, R., *et al.*, *Cell Colonization and Infection Thread Formation em Sugar Cane Roots by *Acetobacter diazotrophicus**. *Soil Biol. Biochem.*, **29** (56): p. 965-967. 1997.
32. James, E.K., Olivares, F.L., de Oliveira, A.L., *et al.*, *Further observations on the interaction between sugar cane and *Gluconacetobacter diazotrophicus* under laboratory and greenhouse conditions*. *J Exp Bot*, **52** (357): p. 747-60. 2001.
33. Meletzus, D., Teixeira, K., Perlova, O., *et al.*, *Characterization of genes involved in regulation of nitrogen fixation and ammonium sensing in *Acetobacter diazotrophicus*, an endophyte of sugarcane*. In. *Biological nitrogen fixation for the 21st century*. Kluwer academic publishers, Dordrecht, The Netherlands.: p. 125-126. 1998.

34. Sevilla, M., Meletzus, D., Teixeira, K., *et al.*, *Analysis of nif and regulatory genes in Acetobacter diazotrophicus*. *Soil Biol. Biochem*, **29**: p. 871-874. 1997.
35. Teixeira, K.R., Morgan, T., Meletzus, D., *et al.*, *Identification, sequencing and structural analysis of a nifA-like gene of Acetobacter diazotrophicus*. *An Acad Bras Cienc*, **71** (3 Pt 2): p. 521-30. 1999.
36. Teixeira, K.R.S., Wulling, M., Morgan, T., *et al.*, *Molecular analysis of the chromosomal region encoding the nifA and nifB genes of Acetobacter diazotrophicus*. *FEMS Microbiol. Lett.*, **71**: p. 521-530. 1999.
37. Kim, J. e Rees, D., *Nitrogenase and biological nitrogen fixation*. *Biochemistry*, **33**: p. 389-397. 1994.
38. Carvalho, E., *Avaliação Agronômica da Disponibilização de Nitrogênio à Cultura de Feijão sob Sistema de Semeadura Direta*. 2002: Piracicaba. p. 63.
39. Morgante, P. *Fixação Biológica e Assimilação de Nitrogênio*. 1996 [cited 2004; Available from: <http://www.ciagri.usp.br/~lazaropp/FisioVegGrad/>].
40. UNICA, *Potencial de Co-geração com resíduos da cana de açúcar: sua compatibilidade com o modelo atual*. 2002, UNCA: São Paulo.
41. Bulen, W.A. e LeComte, J.R., *Nitrogenase complex and its components*. *Methods Enzymol*, **24**: p. 456-70. 1972.
42. Hales, B.J., *Alternative nitrogenase*. *Adv Inorg Biochem*, **8**: p. 165-98. 1990.
43. Ribbe, M., Gadkari, D. e Meyer, O., *N₂ fixation by Streptomyces thermoautotrophicus involves a molybdenum-dinitrogenase and a manganese-superoxide oxidoreductase that couple N₂ reduction to the oxidation of superoxide produced from O₂ by a molybdenum-CO dehydrogenase*. *J Biol Chem*, **272** (42): p. 26627-33. 1997.
44. Dean, D.R., Bolin, J.T. e Zheng, L., *Nitrogenase metalloclusters: structures, organization, and synthesis*. *J Bacteriol*, **175** (21): p. 6737-44. 1993.
45. Chan, J.M., Ryle, M.J. e Seefeldt, L.C., *Evidence that MgATP accelerates primary electron transfer in a Clostridium pasteurianum Fe protein-Azotobacter vinelandii MoFe protein nitrogenase tight complex*. *J Biol Chem*, **274** (25): p. 17593-8. 1999.
46. Zheng, L. e Dean, D.R., *Catalytic formation of a nitrogenase iron-sulfur cluster*. *J Biol Chem*, **269** (29): p. 18723-6. 1994.
47. Peters, J.W., Fisher, K., Newton, W.E., *et al.*, *Involvement of the P cluster in intramolecular electron transfer within the nitrogenase MoFe protein*. *J Biol Chem*, **270** (45): p. 27007-13. 1995.
48. Duyvis, M.G., Wassink, H. e Haaker, H., *Nitrogenase of Azotobacter vinelandii: kinetic analysis of the Fe protein redox cycle*. *Biochemistry*, **37** (50): p. 17345-54. 1998.
49. Shen, J., Dean, D.R. e Newton, W.E., *Evidence for multiple substrate-reduction sites and distinct inhibitor-binding sites from an altered Azotobacter vinelandii nitrogenase MoFe protein*. *Biochemistry*, **36** (16): p. 4884-94. 1997.
50. Lee, S., Reth, A., Meletzus, D., *et al.*, *Characterization of a major cluster of nif, fix, and associated genes in a sugarcane endophyte, Acetobacter diazotrophicus*. *J Bacteriol*, **182** (24): p. 7088-91. 2000.
51. Fisher, K. e Newton, W.E., *Nitrogenase proteins from Gluconacetobacter diazotrophicus, a sugarcane-colonizing bacterium*. *Biochim Biophys Acta*, **1750** (2): p. 154-65. 2005.
52. Earl, C.D., Ronson, C.W. e Ausubel, F.M., *Genetic and structural analysis of the Rhizobium meliloti fixA, fixB, fixC, and fixX genes*. *J Bacteriol*, **169** (3): p. 1127-36. 1987.
53. Mowbray, S.L., *Bacterial chemoreceptors: recent progress in structure and function*. *Mol Cells*, **9** (2): p. 115-8. 1999.

54. Neubauer, H., Pantel, I., Lindgren, P.E., *et al.*, *Characterization of the molybdate transport system ModABC of Staphylococcus carnosus*. Arch Microbiol, **172** (2): p. 109-15. 1999.
55. Reis, V.M. e Dobereiner, J., *Effect of high sugar concentration on nitrogenase activity of Acetobacter diazotrophicus*. Arch Microbiol, **171** (1): p. 13-8. 1998.
56. Robson, R.L. e Postgate, J.R., *Oxygen and hydrogen in biological nitrogen fixation*. Annu Rev Microbiol, **34**: p. 183-207. 1980.
57. Flores-Encarnacion, M., Contreras-Zentella, M., Soto-Urdua, L., *et al.*, *The respiratory system and diazotrophic activity of Acetobacter diazotrophicus PAL5*. J Bacteriol, **181** (22): p. 6987-95. 1999.
58. Dong, Z., Zelmer, C.D., Canny, M.J., *et al.*, *Evidence for protection of nitrogenase from O(2) by colony structure in the aerobic diazotroph Gluconacetobacter diazotrophicus*. Microbiology, **148** (Pt 8): p. 2293-8. 2002.
59. Pan, B. e Vessey, J.K., *Response of the endophytic diazotroph Gluconacetobacter diazotrophicus on solid media to changes in atmospheric partial O(2) pressure*. Appl Environ Microbiol, **67** (10): p. 4694-700. 2001.
60. Ureta, A. e Nordlund, S., *Evidence for conformational protection of nitrogenase against oxygen in Gluconacetobacter diazotrophicus by a putative FeSII protein*. J Bacteriol, **184** (20): p. 5805-9. 2002.
61. Ureta, A. e Nordlund, S., *Glutamine synthetase from Acetobacter diazotrophicus: properties and regulation*. FEMS Microbiol Lett, **202** (2): p. 177-80. 2001.
62. Perlova, O., Ureta, A., Nordlund, S., *et al.*, *Identification of three genes encoding P(II)-like proteins in Gluconacetobacter diazotrophicus: studies of their role(s) in the control of nitrogen fixation*. J Bacteriol, **185** (19): p. 5854-61. 2003.
63. Sevilla, M., Burris, R.H., Gunapala, N., *et al.*, *Comparison of benefit to sugarcane plant growth and 15N2 incorporation following inoculation of sterile plants with Acetobacter diazotrophicus wild-type and Nif- mutants strains*. Mol Plant Microbe Interact, **14** (3): p. 358-66. 2001.
64. Munoz-Rojas, J. e Caballero-Mellado, J., *Population dynamics of Gluconacetobacter diazotrophicus in sugarcane cultivars and its effect on plant growth*. Microb Ecol, **46** (4): p. 454-64. 2003.
65. Piñon, D., Casas, M., Blanch, M., *et al.*, *Gluconacetobacter diazotrophicus, a sugar cane endosymbiont, produces a bacteriocin against Xanthomonas albilineans, a sugar cane pathogen*. Research in Microbiology, **153**: p. 345-351. 2002.
66. Lopes, S., Damann, K., Hoy, J., *et al.*, *Infectivity titration for assesing resistance to leaf scald among sugar cane cultivars*. Plant Dis, **85**: p. 592-596. 2001.
67. Bastian, F., Cohen, A., Piccoli, P., *et al.*, *Production of indole-3-acetic acid and gibberellins A(1) and A(3) by Gluconacetobacter diazotrophicus and Herbaspirillum seropedicae in chemically-defined culture media*. Plant Growth Regul., **24** (1): p. 7-11. 1998.
68. Saravanan, V.S., Madhaiyan, M. e Thangaraju, M., *Solubilization of zinc compounds by the diazotrophic, plant growth promoting bacterium Gluconacetobacter diazotrophicus*. Chemosphere, **66** (9): p. 1794-8. 2007.
69. MRE. Agricultura. 2003 [cited].
70. CNA, *Agronegócio Projeta PIB de R\$458 bi para 2003*. Indicadores Rurais, (50): p. 1-6. 2003.
71. Site do Ministério da Agricultura, P.e.A., <http://www.agricultura.gov.br>. 2005.
72. Coutinho, H., *Biodiversidade: Perspectivas e Oportunidades Tecnológicas*, in *Diversidade Microbiana e Agricultura Sustentável*. 2001, EMBRAPA: <http://www.bdt.fat.org.br/publicacoes>.

73. Boddey, R.M., Urquiaga, S., Reis, V., *et al.*, *Biological nitrogen fixation associated with sugar cane*. Plant and Soil, **137**: p. 111-117. 1991.
74. Cojho, E.H., Reis, V.M., Schenberg, A.C., *et al.*, *Interactions of Acetobacter diazotrophicus with an amylolytic yeast in nitrogen-free batch culture*. Microbiol. Lett., **106**: p. 23-31. 1993.
75. Clarke, S.C., *Pyrosequencing: nucleotide sequencing technology with bacterial genotyping applications*. Expert Rev Mol Diagn, **5** (6): p. 947-53. 2005.
76. Marsh, S., *Pyrosequencing(R) applications*. Methods Mol Biol, **373**: p. 15-24. 2006.
77. Venter, J.C., Adams, M.D., Myers, E.W., *et al.*, *The sequence of the human genome*. Science, **291** (5507): p. 1304-51. 2001.
78. Consortium, I.H.G., *Finishing the euchromatic sequence of the human genome*. Nature, **431** (7011): p. 931-45. 2004.
79. Prust, C., Hoffmeister, M., Liesegang, H., *et al.*, *Complete genome sequence of the acetic acid bacterium Gluconobacter oxydans*. Nat Biotechnol, **23** (2): p. 195-200. 2005.
80. Gupta, A., Singh, V.K., Qazi, G.N., *et al.*, *Gluconobacter oxydans: its biotechnological applications*. J Mol Microbiol Biotechnol, **3** (3): p. 445-56. 2001.
81. Deppenmeier, U., Hoffmeister, M. e Prust, C., *Biochemistry and biotechnological applications of Gluconobacter strains*. Appl Microbiol Biotechnol, **60** (3): p. 233-42. 2002.
82. Wood, D.W., Setubal, J.C., Kaul, R., *et al.*, *The genome of the natural genetic engineer Agrobacterium tumefaciens C58*. Science, **294** (5550): p. 2317-23. 2001.
83. Gonzalez, V., Santamaria, R.I., Bustos, P., *et al.*, *The partitioned Rhizobium etli genome: genetic and metabolic redundancy in seven interacting replicons*. Proc Natl Acad Sci U S A, **103** (10): p. 3834-9. 2006.
84. Kaneko, T., Nakamura, Y., Sato, S., *et al.*, *Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti*. DNA Res, **7** (6): p. 331-8. 2000.
85. Lemos, M., *Workflow para bioinformática*. Tese de doutorado. 2004.
86. Santos, M.B.Q., *Montagem e finalização do genoma de Gluconacetobacter diazotrophicus PAL5*. Dissertação de mestrado. 2005.
87. Wilkins, M.R., Sanchez, J.C., Gooley, A.A., *et al.*, *Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it*. Biotechnol Genet Eng Rev, **13**: p. 19-50. 1996.
88. Celis, J.E., Kruhoffer, M., Gromova, I., *et al.*, *Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics*. FEBS Lett, **480** (1): p. 2-16. 2000.
89. Gromov, P.S. e Celis, J.E., *[From genomics to proteomics]*. Mol Biol (Mosk), **34** (4): p. 597-611. 2000.
90. Wilkins, M.R., Sanchez, J.C., Williams, K.L., *et al.*, *Current challenges and future applications for protein maps and post-translational vector maps in proteome projects*. Electrophoresis, **17** (5): p. 830-8. 1996.
91. Coelho, A., de Oliveira Santos, E., Faria, M.L., *et al.*, *A proteome reference map for Vibrio cholerae El Tor*. Proteomics, **4** (5): p. 1491-504. 2004.
92. Wilkins, M.R., Gasteiger, E., Gooley, A.A., *et al.*, *High-throughput mass spectrometric discovery of protein post-translational modifications*. J Mol Biol, **289** (3): p. 645-57. 1999.
93. Rappsilber, J. e Mann, M., *What does it mean to identify a protein in proteomics?* Trends Biochem Sci, **27** (2): p. 74-8. 2002.
94. Pandey, A. e Mann, M., *Proteomics to study genes and genomes*. Nature, **405** (6788): p. 837-46. 2000.

95. Bjellqvist, B., Ek, K., Righetti, P.G., *et al.*, *Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications*. J Biochem Biophys Methods, **6** (4): p. 317-39. 1982.
96. Gorg, A., Obermaier, C., Boguth, G., *et al.*, *Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures*. Electrophoresis, **20** (4-5): p. 712-7. 1999.
97. Gorg, A., Obermaier, C., Boguth, G., *et al.*, *The current state of two-dimensional electrophoresis with immobilized pH gradients*. Electrophoresis, **21** (6): p. 1037-53. 2000.
98. Kuster, B., Mortensen, P., Andersen, J.S., *et al.*, *Mass spectrometry allows direct identification of proteins in large genomes*. Proteomics, **1** (5): p. 641-50. 2001.
99. Wilkins, M.R., Gasteiger, E., Tonella, L., *et al.*, *Protein identification with N and C-terminal sequence tags in proteome projects*. J Mol Biol, **278** (3): p. 599-608. 1998.
100. Patterson, S.D., *Proteomics: evolution of the technology*. Biotechniques, **35** (3): p. 440-4. 2003.
101. Sivakumar, A., *2D gels and bioinformatics--an eye to the future*. In Silico Biol, **2** (4): p. 507-10. 2002.
102. Nyman, T.A., *The role of mass spectrometry in proteome studies*. Biomol Eng, **18** (5): p. 221-7. 2001.
103. Mann, M., Hendrickson, R.C. e Pandey, A., *Analysis of proteins and proteomes by mass spectrometry*. Annu Rev Biochem, **70**: p. 437-73. 2001.
104. Bisch, P.M., *Genômica Funcional: Proteômica*. In: Luís Mir. (Org.). Genômica. 1 ed. São Paulo: Editora Atheneu., **1**: p. 139-162. 2004.
105. Larmann, J.P., Jr., Lemmo, A.V., Moore, A.W., Jr., *et al.*, *Two-dimensional separations of peptides and proteins by comprehensive liquid chromatography-capillary electrophoresis*. Electrophoresis, **14** (5-6): p. 439-47. 1993.
106. Moore, A.W., Jr. e Jorgenson, J.W., *Comprehensive three-dimensional separation of peptides using size exclusion chromatography/reversed phase liquid chromatography/optically gated capillary zone electrophoresis*. Anal Chem, **67** (19): p. 3456-63. 1995.
107. Moore, A.W., Jr., Larmann, J.P., Jr., Lemmo, A.V., *et al.*, *Two-dimensional liquid chromatography-capillary electrophoresis techniques for analysis of proteins and peptides*. Methods Enzymol, **270**: p. 401-19. 1996.
108. Kislinger, T., Gramolini, A.O., MacLennan, D.H., *et al.*, *Multidimensional protein identification technology (MudPIT): technical overview of a profiling method optimized for the comprehensive proteomic investigation of normal and diseased heart tissue*. J Am Soc Mass Spectrom, **16** (8): p. 1207-20. 2005.
109. Chen, E.I., Hewel, J., Felding-Habermann, B., *et al.*, *Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT)*. Mol Cell Proteomics, **5** (1): p. 53-6. 2006.
110. Schirmer, E.C., Yates, J.R. e Gerace, L., *MudPIT: A powerful proteomics tool for discovery, limited primarily by our imagination*. Discovery Medicine, **3**: p. 38-39. 2003.
111. Jungblut, P.R., Schaible, U.E., Mollenkopf, H.J., *et al.*, *Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens*. Mol Microbiol, **33** (6): p. 1103-17. 1999.

112. Wasinger, V.C., Cordwell, S.J., Cerpa-Poljak, A., *et al.*, *Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium*. *Electrophoresis*, **16** (7): p. 1090-4. 1995.
113. von Kruger, W.M., Lery, L.M., Soares, M.R., *et al.*, *The phosphate-starvation response in Vibrio cholerae O1 and phoB mutant under proteomic analysis: disclosing functions involved in adaptation, survival and virulence*. *Proteomics*, **6** (5): p. 1495-511. 2006.
114. Vohradsky, J. e Thompson, C.J., *Systems level analysis of protein synthesis patterns associated with bacterial growth and metabolic transitions*. *Proteomics*, **6** (3): p. 785-93. 2006.
115. Silberbach, M. e Burkovski, A., *Application of global analysis techniques to Corynebacterium glutamicum: new insights into nitrogen regulation*. *J Biotechnol*, **126** (1): p. 101-10. 2006.
116. Silberbach, M., Schafer, M., Huser, A.T., *et al.*, *Adaptation of Corynebacterium glutamicum to ammonium limitation: a global analysis using transcriptome and proteome techniques*. *Appl Environ Microbiol*, **71** (5): p. 2391-402. 2005.
117. Anderson, D.C., Campbell, E.L. e Meeks, J.C., *A soluble 3D LC/MS/MS proteome of the filamentous cyanobacterium Nostoc punctiforme*. *J Proteome Res*, **5** (11): p. 3096-104. 2006.
118. Chen, H., Teplitski, M., Robinson, J.B., *et al.*, *Proteomic analysis of wild-type Sinorhizobium meliloti responses to N-acyl homoserine lactone quorum-sensing signals and the transition to stationary phase*. *J Bacteriol*, **185** (17): p. 5029-36. 2003.
119. Guerreiro, N., Djordjevic, M.A. e Rolfe, B.G., *Proteome analysis of the model microsymbiont Sinorhizobium meliloti: isolation and characterisation of novel proteins*. *Electrophoresis*, **20** (4-5): p. 818-25. 1999.
120. Djordjevic, M.A., Chen, H.C., Natera, S., *et al.*, *A global analysis of protein expression profiles in Sinorhizobium meliloti: discovery of new genes for nodule occupancy and stress adaptation*. *Mol Plant Microbe Interact*, **16** (6): p. 508-24. 2003.
121. Munchbach, M., Dainese, P., Staudenmann, W., *et al.*, *Proteome analysis of heat shock protein expression in Bradyrhizobium japonicum*. *Eur J Biochem*, **264** (1): p. 39-48. 1999.
122. Rosen, R., Buttner, K., Becher, D., *et al.*, *Heat shock proteome of Agrobacterium tumefaciens: evidence for new control systems*. *J Bacteriol*, **184** (6): p. 1772-8. 2002.
123. Steiner, P. e Sauer, U., *Proteins induced during adaptation of Acetobacter aceti to high acetate concentrations*. *Appl Environ Microbiol*, **67** (12): p. 5474-81. 2001.
124. DelVecchio, V.G., Wagner, M.A., Eschenbrenner, M., *et al.*, *Brucella proteomes--a review*. *Vet Microbiol*, **90** (1-4): p. 593-603. 2002.
125. Sinchaikul, S., Sookkheo, B., Topanuruk, S., *et al.*, *Bioinformatics, functional genomics, and proteomics study of Bacillus sp.* *J Chromatogr B Analyt Technol Biomed Life Sci*, **771** (1-2): p. 261-87. 2002.
126. Mihoub, F., Mistou, M.Y., Guillot, A., *et al.*, *Cold adaptation of Escherichia coli: microbiological and proteomic approaches*. *Int J Food Microbiol*, **89** (2-3): p. 171-84. 2003.
127. Hecker, M., Engelmann, S. e Cordwell, S.J., *Proteomics of Staphylococcus aureus--current state and future challenges*. *J Chromatogr B Analyt Technol Biomed Life Sci*, **787** (1): p. 179-95. 2003.
128. Guina, T., Wu, M., Miller, S.I., *et al.*, *Proteomic analysis of Pseudomonas aeruginosa grown under magnesium limitation*. *J Am Soc Mass Spectrom*, **14** (7): p. 742-51. 2003.

129. Len, A.C., Cordwell, S.J., Harty, D.W., *et al.*, *Cellular and extracellular proteome analysis of Streptococcus mutans grown in a chemostat*. Proteomics, **3** (5): p. 627-46. 2003.
130. Pizzatti, L., Sa, L.A., de Souza, J.M., *et al.*, *Altered protein profile in chronic myeloid leukemia chronic phase identified by a comparative proteomic study*. Biochim Biophys Acta, **1764** (5): p. 929-42. 2006.
131. Lery, L.M.S., *Análise de proteomas diferenciais da bactéria endofítica Gluconacetobacter diazotrophicus em condições de fixação e não fixação de nitrogênio*. Monografia de final de curso de graduação em Ciências Biológicas - Modalidade Médica. 2005.
132. Kruger, N.J., *The Bradford method for protein quantitation*. Methods Mol Biol, **32**: p. 9-15. 1994.
133. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, **227** (5259): p. 680-5. 1970.
134. Berkelman, T. e Stenstedt, T., *2-D Electrophoresis using Immobilized pH Gradients: Principles and Methods*. Piscataway, NJ. 1998.
135. Clauser, K.R., Baker, P. e Burlingame, A.L., *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching*. Anal Chem, **71** (14): p. 2871-82. 1999.
136. Perkins, D.N., Pappin, D.J., Creasy, D.M., *et al.*, *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, **20** (18): p. 3551-67. 1999.
137. Gobom, J., Nordhoff, E., Mirgorodskaya, E., *et al.*, *Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry*. J Mass Spectrom, **34** (2): p. 105-16. 1999.
138. Delcher, A.L., Harmon, D., Kasif, S., *et al.*, *Improved microbial gene identification with GLIMMER*. Nucleic Acids Res, **27** (23): p. 4636-41. 1999.
139. Altschul, S.F., Gish, W., Miller, W., *et al.*, *Basic local alignment search tool*. J Mol Biol, **215** (3): p. 403-10. 1990.
140. Altschul, S.F., Madden, T.L., Schaffer, A.A., *et al.*, *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. Nucleic Acids Res, **25** (17): p. 3389-402. 1997.
141. Bateman, A., Coin, L., Durbin, R., *et al.*, *The Pfam protein families database*. Nucleic Acids Res, **32** (Database issue): p. D138-41. 2004.
142. Tatusov, R.L., Koonin, E.V. e Lipman, D.J., *A genomic perspective on protein families*. Science, **278** (5338): p. 631-7. 1997.
143. Kanehisa, M., Goto, S., Hattori, M., *et al.*, *From genomics to chemical genomics: new developments in KEGG*. Nucleic Acids Res, **34** (Database issue): p. D354-7. 2006.
144. Sonnhammer, E.L., von Heijne, G. e Krogh, A., *A hidden Markov model for predicting transmembrane helices in protein sequences*. Proc Int Conf Intell Syst Mol Biol, **6**: p. 175-82. 1998.
145. Bendtsen, J.D., Nielsen, H., von Heijne, G., *et al.*, *Improved prediction of signal peptides: SignalP 3.0*. J Mol Biol, **340** (4): p. 783-95. 2004.
146. Rutherford, K., Parkhill, J., Crook, J., *et al.*, *Artemis: sequence visualization and annotation*. Bioinformatics, **16** (10): p. 944-5. 2000.
147. Len, A.C., Harty, D.W. e Jacques, N.A., *Proteome analysis of Streptococcus mutans metabolic phenotype during acid tolerance*. Microbiology, **150** (Pt 5): p. 1353-66. 2004.

148. Hurek, T., Van Montagu, M., Kellenberger, E., *et al.*, *Induction of complex intracytoplasmic membranes related to nitrogen fixation in Azoarcus sp. BH72*. Mol Microbiol, **18** (2): p. 225-36. 1995.
149. Karg, T. e Reinhold-Hurek, B., *Global changes in protein composition of N₂-fixing-Azoarcus sp. strain BH72 upon diazosome formation*. J Bacteriol, **178** (19): p. 5748-54. 1996.
150. Feng, L., Roughley, R.J. e Copeland, L., *Morphological changes of rhizobia in peat cultures*. Appl Environ Microbiol, **68** (3): p. 1064-70. 2002.
151. Twells, R., Thomson, N., Parkhill, J., *et al.*, *Working with pathogen genomes*. The first Joint Pasteur Institute / Wellcome Trust course on genomics in South America. 2006.
152. Nachin, L., El Hassouni, M., Loiseau, L., *et al.*, *SoxR-dependent response to oxidative stress and virulence of Erwinia chrysanthemi: the key role of SufC, an orphan ABC ATPase*. Mol Microbiol, **39** (4): p. 960-72. 2001.
153. Murat, D., Bance, P., Callebaut, I., *et al.*, *ATP hydrolysis is essential for the function of the Uup ATP-binding cassette ATPase in precise excision of transposons*. J Biol Chem, **281** (10): p. 6850-9. 2006.
154. Abergel, C., Bouveret, E., Claverie, J.M., *et al.*, *Structure of the Escherichia coli TolB protein determined by MAD methods at 1.95 Å resolution*. Structure, **7** (10): p. 1291-300. 1999.
155. Simpson, W., Olczak, T. e Genco, C.A., *Characterization and expression of HmuR, a TonB-dependent hemoglobin receptor of Porphyromonas gingivalis*. J Bacteriol, **182** (20): p. 5737-48. 2000.
156. Driessen, A.J., *SecB, a molecular chaperone with two faces*. Trends Microbiol, **9** (5): p. 193-6. 2001.
157. Gilson, L., Mahanty, H.K. e Kolter, R., *Genetic analysis of an MDR-like export system: the secretion of colicin V*. Embo J, **9** (12): p. 3875-84. 1990.
158. Pucci, M.J., Discotto, L.F. e Dougherty, T.J., *Cloning and identification of the Escherichia coli murB DNA sequence, which encodes UDP-N-acetylenolpyruvoylglucosamine reductase*. J Bacteriol, **174** (5): p. 1690-3. 1992.
159. Malinverni, J.C., Werner, J., Kim, S., *et al.*, *YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli*. Mol Microbiol, **61** (1): p. 151-64. 2006.
160. Werner, J. e Misra, R., *YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of Escherichia coli*. Mol Microbiol, **57** (5): p. 1450-9. 2005.
161. Rees, D.C. e Howard, J.B., *Nitrogenase: standing at the crossroads*. Curr Opin Chem Biol, **4** (5): p. 559-66. 2000.
162. Kim, J. e Rees, D.C., *Nitrogenase and biological nitrogen fixation*. Biochemistry, **33** (2): p. 389-97. 1994.
163. Pawlowski, K., Klosse, U. e de Bruijn, F.J., *Characterization of a novel Azorhizobium caulinodans ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism*. Mol Gen Genet, **231** (1): p. 124-38. 1991.
164. Pinheiro, I., *Paraquat poisoning – international bibliographic review*. SPMI, **9** (4). 2002.
165. Hartl, F.U., *Molecular chaperones in cellular protein folding*. Nature, **381** (6583): p. 571-9. 1996.
166. Bukau, *Assisted protein folding*. In: Molecular chaperones and folding catalysts – regulation, cellular function and mechanisms, **Harwood academic publishers**. 1999.
167. Hendrick, J.P. e Hartl, F.U., *Molecular chaperone functions of heat-shock proteins*. Annu Rev Biochem, **62**: p. 349-84. 1993.

168. Lazar, S.W. e Kolter, R., *SurA assists the folding of Escherichia coli outer membrane proteins*. J Bacteriol, **178** (6): p. 1770-3. 1996.
169. Bardwell, J.C. e Craig, E.A., *Ancient heat shock gene is dispensable*. J Bacteriol, **170** (7): p. 2977-83. 1988.
170. Tanaka, N. e Nakamoto, H., *HtpG is essential for the thermal stress management in cyanobacteria*. FEBS Lett, **458** (2): p. 117-23. 1999.
171. Weber, M.H. e Marahiel, M.A., *Bacterial cold shock responses*. Sci Prog, **86** (Pt 1-2): p. 9-75. 2003.
172. Ermolenko, D.N. e Makhatadze, G.I., *Bacterial cold-shock proteins*. Cell Mol Life Sci, **59** (11): p. 1902-13. 2002.
173. Schwartz, C.E., Krall, J., Norton, L., et al., *Catalase and superoxide dismutase in Escherichia coli*. J Biol Chem, **258** (10): p. 6277-81. 1983.
174. Wood, Z.A., Schroder, E., Robin Harris, J., et al., *Structure, mechanism and regulation of peroxiredoxins*. Trends Biochem Sci, **28** (1): p. 32-40. 2003.
175. Santos, J.M., Lobo, M., Matos, A.P., et al., *The gene bolA regulates dacA (PBP5), dacC (PBP6) and ampC (AmpC), promoting normal morphology in Escherichia coli*. Mol Microbiol, **45** (6): p. 1729-40. 2002.
176. Andersson, S.G., Zomorodipour, A., Andersson, J.O., et al., *The genome sequence of Rickettsia prowazekii and the origin of mitochondria*. Nature, **396** (6707): p. 133-40. 1998.
177. Keller, M., Roxlau, A., Weng, W.M., et al., *Molecular analysis of the Rhizobium meliloti mucR gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan*. Mol Plant Microbe Interact, **8** (2): p. 267-77. 1995.
178. Cheng, H.P. e Walker, G.C., *Succinoglycan production by Rhizobium meliloti is regulated through the ExoS-ChvI two-component regulatory system*. J Bacteriol, **180** (1): p. 20-6. 1998.
179. Baev, N., Endre, G., Petrovics, G., et al., *Six nodulation genes of nod box locus 4 in Rhizobium meliloti are involved in nodulation signal production: nodM codes for D-glucosamine synthetase*. Mol Gen Genet, **228** (1-2): p. 113-24. 1991.
180. Baev, N., Schultze, M., Barlier, I., et al., *Rhizobium nodM and nodN genes are common nod genes: nodM encodes functions for efficiency of nod signal production and bacteroid maturation*. J Bacteriol, **174** (23): p. 7555-65. 1992.
181. Preisig, O., Zufferey, R. e Hennecke, H., *The Bradyrhizobium japonicum fixGHIS genes are required for the formation of the high-affinity cbb3-type cytochrome oxidase*. Arch Microbiol, **165** (5): p. 297-305. 1996.
182. Miyasaka, N.R., Thuler, D.S., Floh, E.I., et al., *During stationary phase, Beijerinckia derxii shows nitrogenase activity concomitant with the release and accumulation of nitrogenated substances*. Microbiol Res, **158** (4): p. 309-15. 2003.
183. Riley, M., *Functions of the gene products of Escherichia coli*. Microbiol Rev, **57** (4): p. 862-952. 1993.
184. Roberts, R.J. e Macelis, D., *REBASE--restriction enzymes and methylases*. Nucleic Acids Res, **24** (1): p. 223-35. 1996.
185. Treuner-Lange, A., Kuhn, A. e Durre, P., *The kdp system of Clostridium acetobutylicum: cloning, sequencing, and transcriptional regulation in response to potassium concentration*. J Bacteriol, **179** (14): p. 4501-12. 1997.
186. Walderhaug, M.O., Polarek, J.W., Voelkner, P., et al., *KdpD and KdpE, proteins that control expression of the kdpABC operon, are members of the two-component sensor-effector class of regulators*. J Bacteriol, **174** (7): p. 2152-9. 1992.
187. Kroll, R.G. e Booth, I.R., *The relationship between intracellular pH, the pH gradient and potassium transport in Escherichia coli*. Biochem J, **216** (3): p. 709-16. 1983.

188. Slonczewski, J.L., Macnab, R.M., Alger, J.R., *et al.*, *Effects of pH and repellent tactic stimuli on protein methylation levels in Escherichia coli*. J Bacteriol, **152** (1): p. 384-99. 1982.
189. Bakker, E.P. e Mangerich, W.E., *Interconversion of components of the bacterial proton motive force by electrogenic potassium transport*. J Bacteriol, **147** (3): p. 820-6. 1981.
190. Altendorf, K., Gassel, M., Puppe, W., *et al.*, *Structure and function of the Kdp-ATPase of Escherichia coli*. Acta Physiol Scand Suppl, **643**: p. 137-46. 1998.
191. Fendler, K., Drose, S., Epstein, W., *et al.*, *The Kdp-ATPase of Escherichia coli mediates an ATP-dependent, K⁺-independent electrogenic partial reaction*. Biochemistry, **38** (6): p. 1850-6. 1999.
192. Borghese, R., Crimi, M., Fava, L., *et al.*, *The ATP synthase atpHAGDC (F1) operon from Rhodobacter capsulatus*. J Bacteriol, **180** (2): p. 416-21. 1998.
193. Tran, S.L. e Cook, G.M., *The F1Fo-ATP synthase of Mycobacterium smegmatis is essential for growth*. J Bacteriol, **187** (14): p. 5023-8. 2005.
194. Borghese, R., Turina, P., Lambertini, L., *et al.*, *The atpIBEXF operon coding for the F0 sector of the ATP synthase from the purple nonsulfur photosynthetic bacterium Rhodobacter capsulatus*. Arch Microbiol, **170** (5): p. 385-8. 1998.
195. Sheng, J. e Marquis, R.E., *Enhanced acid resistance of oral streptococci at lethal pH values associated with acid-tolerant catabolism and with ATP synthase activity*. FEMS Microbiol Lett, **262** (1): p. 93-8. 2006.
196. Belli, W.A. e Marquis, R.E., *Adaptation of Streptococcus mutans and Enterococcus hirae to acid stress in continuous culture*. Appl Environ Microbiol, **57** (4): p. 1134-8. 1991.
197. Rees, D.C., *Great metalloclusters in enzymology*. Annu Rev Biochem, **71**: p. 221-46. 2002.
198. Hille, R., *Molybdenum and tungsten in biology*. Trends Biochem Sci, **27** (7): p. 360-7. 2002.
199. Che, P., Weaver, L.M., Wurtele, E.S., *et al.*, *The role of biotin in regulating 3-methylcrotonyl-coenzyme a carboxylase expression in Arabidopsis*. Plant Physiol, **131** (3): p. 1479-86. 2003.
200. Schneider, G. e Lindqvist, Y., *Structural enzymology of biotin biosynthesis*. FEBS Lett, **495** (1-2): p. 7-11. 2001.
201. Streit, W.R. e Entcheva, P., *Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production*. Appl Microbiol Biotechnol, **61** (1): p. 21-31. 2003.
202. Galibert, F., Finan, T.M., Long, S.R., *et al.*, *The composite genome of the legume symbiont Sinorhizobium meliloti*. Science, **293** (5530): p. 668-72. 2001.
203. Kaneko, T., Nakamura, Y., Sato, S., *et al.*, *Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti (supplement)*. DNA Res, **7** (6): p. 381-406. 2000.
204. Rodrigues Neto, J., Malavolta Jr., V.A. e Victor, O., *Meio simples para isolamento e cultivo de Xanthomonas campestris pv. citri Tipo B*. Sum. Phytol., **2**: p. 16. 1986.
205. Sundaram, T.K., Chell, R.M. e Wilkinson, A.E., *Monomeric malate synthase from a thermophilic Bacillus. Molecular and kinetic characteristics*. Arch Biochem Biophys, **199** (2): p. 515-25. 1980.
206. Johnson, G.V., Evans, H.J. e Ching, T., *Enzymes of the Glyoxylate Cycle in Rhizobia and Nodules of Legumes*. Plant Physiol, **41** (8): p. 1330-1336. 1966.
207. Stovall, I. e Cole, M., *Organic Acid Metabolism by Isolated Rhizobium japonicum Bacteroids*. Plant Physiol, **61** (5): p. 787-790. 1978.

208. Karr, D.B., Waters, J.K., Suzuki, F., *et al.*, *Enzymes of the Poly-beta-Hydroxybutyrate and Citric Acid Cycles of Rhizobium japonicum Bacteroids*. Plant Physiol, **75** (4): p. 1158-1162. 1984.
209. Green, L.S., Karr, D.B. e Emerich, D.W., *Isocitrate dehydrogenase and glyoxylate cycle enzyme activities in Bradyrhizobium japonicum under various growth conditions*. Arch Microbiol, **169** (5): p. 445-51. 1998.
210. Mandal, N.C. e Chakrabarty, P.K., *Enzymes of carbohydrate metabolism in fast-growing Rhizobium grown on hexoses or succinate*. Indian J Biochem Biophys, **26** (2): p. 120-2. 1989.
211. Kornberg, H.L., Gotto, A.M. e Lund, P., *Effect of growth substrates on isocitratase formation by Pseudomonas ovalis Chester*. Nature, **182** (4647): p. 1430-1. 1958.
212. Gottschalk, G., *Bacterial metabolism*. Springer, New York, Berlin, Heidelberg. 1986.
213. McDermott, T.R. e Graham, P.H., *Bradyrhizobium japonicum Inoculant Mobility, Nodule Occupancy, and Acetylene Reduction in the Soybean Root System*. Appl Environ Microbiol, **55** (10): p. 2493-2498. 1989.
214. Harwood, C.S. e Parales, R.E., *The beta-ketoadipate pathway and the biology of self-identity*. Annu Rev Microbiol, **50**: p. 553-90. 1996.
215. Harwood, C.S., Nichols, N.N., Kim, M.K., *et al.*, *Identification of the pcaRKF gene cluster from Pseudomonas putida: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate*. J Bacteriol, **176** (21): p. 6479-88. 1994.
216. Lopez-de-Victoria, G. e Lovell, C.R., *Chemotaxis of Azospirillum Species to Aromatic Compounds*. Appl Environ Microbiol, **59** (9): p. 2951-2955. 1993.
217. Sundin, G.W., Shankar, S., Chugani, S.A., *et al.*, *Nucleoside diphosphate kinase from Pseudomonas aeruginosa: characterization of the gene and its role in cellular growth and exopolysaccharide alginate synthesis*. Mol Microbiol, **20** (5): p. 965-79. 1996.
218. Chakrabarty, A.M., *Nucleoside diphosphate kinase: role in bacterial growth, virulence, cell signalling and polysaccharide synthesis*. Mol Microbiol, **28** (5): p. 875-82. 1998.
219. Zaphiropoulos, P.G., *Non-homologous recombination mediated by Thermus aquaticus DNA polymerase I. Evidence supporting a copy choice mechanism*. Nucleic Acids Res, **26** (12): p. 2843-8. 1998.
220. Smith, M.C. e Thorpe, H.M., *Diversity in the serine recombinases*. Mol Microbiol, **44** (2): p. 299-307. 2002.
221. Lawrence, J.G. e Hendrickson, H., *Lateral gene transfer: when will adolescence end?* Mol Microbiol, **50** (3): p. 739-49. 2003.
222. Goenrich, M., Bartoschek, S., Hagemeyer, C.H., *et al.*, *A glutathione-dependent formaldehyde-activating enzyme (Gfa) from Paracoccus denitrificans detected and purified via two-dimensional proton exchange NMR spectroscopy*. J Biol Chem, **277** (5): p. 3069-72. 2002.
223. Neculai, A.M., Neculai, D., Griesinger, C., *et al.*, *A dynamic zinc redox switch*. J Biol Chem, **280** (4): p. 2826-30. 2005.
224. Kiley, P.J. e Beinert, H., *The role of Fe-S proteins in sensing and regulation in bacteria*. Curr Opin Microbiol, **6** (2): p. 181-5. 2003.
225. Yuvaniyama, P., Agar, J.N., Cash, V.L., *et al.*, *NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein*. Proc Natl Acad Sci U S A, **97** (2): p. 599-604. 2000.
226. Tuteja, R., *Type I signal peptidase: an overview*. Arch Biochem Biophys, **441** (2): p. 107-11. 2005.

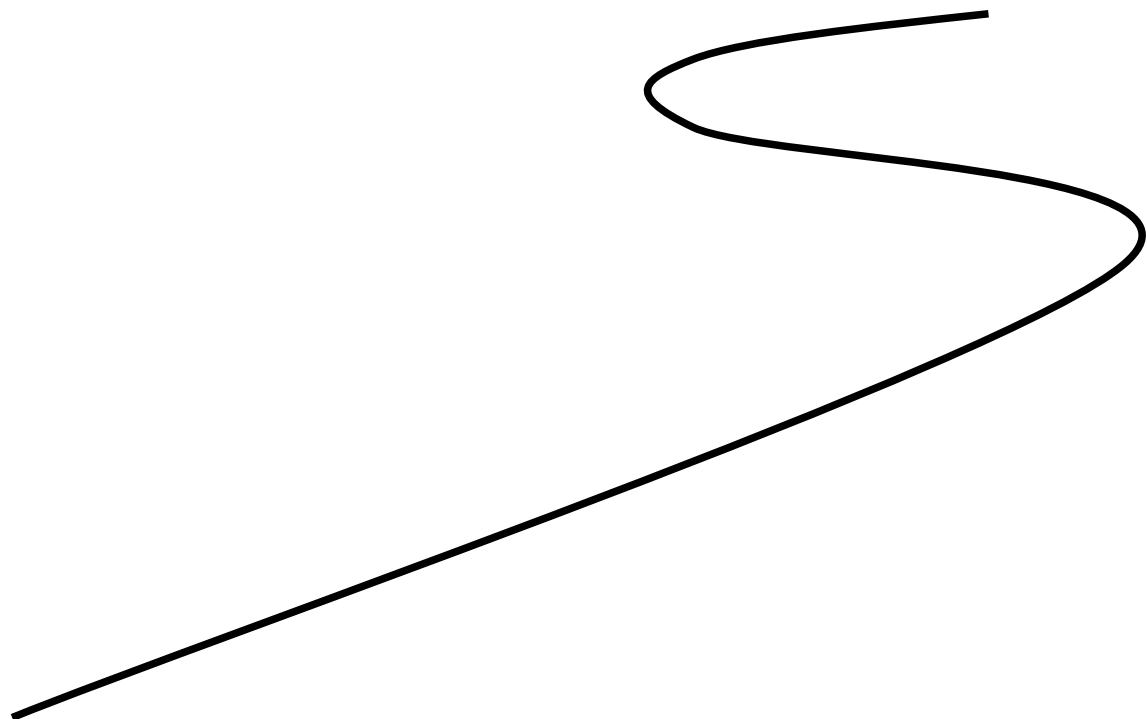
227. Sekowska, A., Kung, H.F. e Danchin, A., *Sulfur metabolism in Escherichia coli and related bacteria: facts and fiction*. J Mol Microbiol Biotechnol, **2** (2): p. 145-77. 2000.
228. Yamaoka, T., Yano, M., Kondo, M., et al., *Feedback inhibition of amidophosphoribosyltransferase regulates the rate of cell growth via purine nucleotide, DNA, and protein syntheses*. J Biol Chem, **276** (24): p. 21285-91. 2001.
229. Keer, J., Smeulders, M.J., Gray, K.M., et al., *Mutants of Mycobacterium smegmatis impaired in stationary-phase survival*. Microbiology, **146** (Pt 9): p. 2209-17. 2000.
230. Liu, Y., Tsinoremas, N.F., Golden, S.S., et al., *Circadian expression of genes involved in the purine biosynthetic pathway of the cyanobacterium Synechococcus sp. strain PCC 7942*. Mol Microbiol, **20** (5): p. 1071-81. 1996.
231. Wu, L.F., Tomich, J.M. e Saier, M.H., Jr., *Structure and evolution of a multidomain multiphosphoryl transfer protein. Nucleotide sequence of the fruB(HI) gene in Rhodobacter capsulatus and comparisons with homologous genes from other organisms*. J Mol Biol, **213** (4): p. 687-703. 1990.
232. Gutknecht, R., Beutler, R., Garcia-Alles, L.F., et al., *The dihydroxyacetone kinase of Escherichia coli utilizes a phosphoprotein instead of ATP as phosphoryl donor*. Embo J, **20** (10): p. 2480-6. 2001.
233. Barabote, R.D. e Saier, M.H., Jr., *Comparative genomic analyses of the bacterial phosphotransferase system*. Microbiol Mol Biol Rev, **69** (4): p. 608-34. 2005.
234. Goodell, E.W. e Higgins, C.F., *Uptake of cell wall peptides by Salmonella typhimurium and Escherichia coli*. J Bacteriol, **169** (8): p. 3861-5. 1987.
235. Abouhamad, W.N. e Manson, M.D., *The dipeptide permease of Escherichia coli closely resembles other bacterial transport systems and shows growth-phase-dependent expression*. Mol Microbiol, **14** (5): p. 1077-92. 1994.
236. Matsuzaki, M., Kiso, Y., Yamamoto, I., et al., *Gene disruption analysis of DppA isolated as a periplasmic molecular chaperone-like protein for folding of dimethyl sulfoxide reductase in Rhodobacter sphaeroides f. sp. denitrificans*. FEMS Microbiol Lett, **193** (2): p. 223-9. 2000.
237. Hanson, M.S., Slaughter, C. e Hansen, E.J., *The hbpA gene of Haemophilus influenzae type b encodes a heme-binding lipoprotein conserved among heme-dependent Haemophilus species*. Infect Immun, **60** (6): p. 2257-66. 1992.
238. Carter, R.A., Yeoman, K.H., Klein, A., et al., *dpp genes of Rhizobium leguminosarum specify uptake of delta-aminolevulinic acid*. Mol Plant Microbe Interact, **15** (1): p. 69-74. 2002.
239. Elliott, T., *Transport of 5-aminolevulinic acid by the dipeptide permease in Salmonella typhimurium*. J Bacteriol, **175** (2): p. 325-31. 1993.
240. Verkamp, E., Backman, V.M., Bjornsson, J.M., et al., *The periplasmic dipeptide permease system transports 5-aminolevulinic acid in Escherichia coli*. J Bacteriol, **175** (5): p. 1452-6. 1993.
241. Lopilato, J.E., Garwin, J.L., Emr, S.D., et al., *D-ribose metabolism in Escherichia coli K-12: genetics, regulation, and transport*. J Bacteriol, **158** (2): p. 665-73. 1984.
242. Wang, R., Prince, J.T. e Marcotte, E.M., *Mass spectrometry of the M. smegmatis proteome: protein expression levels correlate with function, operons, and codon bias*. Genome Res, **15** (8): p. 1118-26. 2005.
243. Sutherland, I.W., *Biosynthesis and composition of gram-negative bacterial extracellular and wall polysaccharides*. Annu Rev Microbiol, **39**: p. 243-70. 1985.
244. Roset, M.S., Ciocchini, A.E., Ugalde, R.A., et al., *Molecular cloning and characterization of cgt, the Brucella abortus cyclic beta-1,2-glucan transporter gene, and its role in virulence*. Infect Immun, **72** (4): p. 2263-71. 2004.

245. Ugalde, R.A., *Intracellular lifestyle of Brucella spp. Common genes with other animal pathogens, plant pathogens, and endosymbionts*. *Microbes Infect*, **1** (14): p. 1211-9. 1999.
246. Briones, G., Inon de Iannino, N., Roset, M., *et al.*, *Brucella abortus cyclic beta-1,2-glucan mutants have reduced virulence in mice and are defective in intracellular replication in HeLa cells*. *Infect Immun*, **69** (7): p. 4528-35. 2001.
247. Hengge-Aronis, R., *Stationary phase gene regulation: what makes an Escherichia coli promoter sigmaS-selective?* *Curr Opin Microbiol*, **5** (6): p. 591-5. 2002.
248. Beyer-Sehlmeyer, G., Kreikemeyer, B., Horster, A., *et al.*, *Analysis of the growth phase-associated transcriptome of Streptococcus pyogenes*. *Int J Med Microbiol*, **295** (3): p. 161-77. 2005.
249. Putzer, H., Laalami, S., Brakhage, A.A., *et al.*, *Aminoacyl-tRNA synthetase gene regulation in Bacillus subtilis: induction, repression and growth-rate regulation*. *Mol Microbiol*, **16** (4): p. 709-18. 1995.
250. Lee, P.C., Bochner, B.R. e Ames, B.N., *AppppA, heat-shock stress, and cell oxidation*. *Proc Natl Acad Sci U S A*, **80** (24): p. 7496-500. 1983.
251. Lee, P.C., Bochner, B.R. e Ames, B.N., *Diadenosine 5',5'''-P1,P4-tetraphosphate and related adenylylated nucleotides in Salmonella typhimurium*. *J Biol Chem*, **258** (11): p. 6827-34. 1983.
252. Putzer, H. e Laalami, S., *Regulation of the Expression of Aminoacyl-tRNA Synthetases and Translation Factors*. *Translation Mechanisms*. 2003.
253. Spreitzer, R.J. e Salvucci, M.E., *Rubisco: structure, regulatory interactions, and possibilities for a better enzyme*. *Annu Rev Plant Biol*, **53**: p. 449-75. 2002.
254. Ashida, H., Danchin, A. e Yokota, A., *Was photosynthetic RuBisCO recruited by acquisitive evolution from RuBisCO-like proteins involved in sulfur metabolism?* *Res Microbiol*, **156** (5-6): p. 611-8. 2005.
255. Ashida, H., Saito, Y., Kojima, C., *et al.*, *A functional link between RuBisCO-like protein of Bacillus and photosynthetic RuBisCO*. *Science*, **302** (5643): p. 286-90. 2003.
256. Hanson, T.E. e Tabita, F.R., *A ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)-like protein from Chlorobium tepidum that is involved with sulfur metabolism and the response to oxidative stress*. *Proc Natl Acad Sci U S A*, **98** (8): p. 4397-402. 2001.
257. Flemming, D., Hellwig, P., Lepper, S., *et al.*, *Catalytic Importance of Acidic Amino Acids on Subunit NuoB of the Escherichia coli NADH:Ubiquinone Oxidoreductase (Complex I)*. *J Biol Chem*, **281** (34): p. 24781-9. 2006.
258. Tao, M., Turk, K., Diez, J., *et al.*, *Crystallization of the NADH-oxidizing domain of the Na⁺-translocating NADH:ubiquinone oxidoreductase from Vibrio cholerae*. *Acta Crystallograph Sect F Struct Biol Cryst Commun*, **62** (Pt 2): p. 110-2. 2006.
259. Aravind, L., Anantharaman, V. e Koonin, E.V., *Monophyly of class I aminoacyl tRNA synthetase, USPA, ETFP, photolyase, and PP-ATPase nucleotide-binding domains: implications for protein evolution in the RNA*. *Proteins*, **48** (1): p. 1-14. 2002.
260. John, M., Schmidt, J., Wieneke, U., *et al.*, *Transmembrane orientation and receptor-like structure of the Rhizobium meliloti common nodulation protein NodC*. *Embo J*, **7** (3): p. 583-588. 1988.
261. Kamst, E., Pilling, J., Raamsdonk, L.M., *et al.*, *Rhizobium nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis*. *J Bacteriol*, **179** (7): p. 2103-8. 1997.
262. Hirsch, A.M., Lum, M.R. e Downie, J.A., *What makes the rhizobia-legume symbiosis so special?* *Plant Physiol*, **127** (4): p. 1484-92. 2001.

263. Buchanan, J.M., *The amidotransferases*. Adv Enzymol Relat Areas Mol Biol, **39**: p. 91-183. 1973.
264. Goffin, C. e Ghuysen, J.M., *Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs*. Microbiol Mol Biol Rev, **62** (4): p. 1079-93. 1998.
265. Den Blaauwen, T., Aarsman, M.E., Vischer, N.O., *et al.*, *Penicillin-binding protein PBP2 of Escherichia coli localizes preferentially in the lateral wall and at mid-cell in comparison with the old cell pole*. Mol Microbiol, **47** (2): p. 539-47. 2003.
266. Scheffers, D.J. e Pinho, M.G., *Bacterial cell wall synthesis: new insights from localization studies*. Microbiol Mol Biol Rev, **69** (4): p. 585-607. 2005.
267. Nanninga, N., *Morphogenesis of Escherichia coli*. Microbiol Mol Biol Rev, **62** (1): p. 110-29. 1998.
268. Lee, H.Y., An, J.H. e Kim, Y.S., *Identification and characterization of a novel transcriptional regulator, MatR, for malonate metabolism in Rhizobium leguminosarum bv. trifolii*. Eur J Biochem, **267** (24): p. 7224-30. 2000.
269. Lee, H.Y. e Kim, Y.S., *Transcription of matR gene in Rhizobium leguminosarum bv. trifolii*. J Biochem Mol Biol Biophys, **6** (4): p. 283-8. 2002.
270. Stumpf, D.K. e Burris, R.H., *Organic Acid Contents of Soybean: Age and Source of Nitrogen*. Plant Physiol, **68** (5): p. 989-991. 1981.
271. Kahn, D., David, M., Domergue, O., *et al.*, *Rhizobium meliloti fixGHI sequence predicts involvement of a specific cation pump in symbiotic nitrogen fixation*. J Bacteriol, **171** (2): p. 929-39. 1989.
272. Neidle, E.L. e Kaplan, S., *Rhodobacter sphaeroides rdxA, a homolog of Rhizobium meliloti fixG, encodes a membrane protein which may bind cytoplasmic [4Fe-4S] clusters*. J Bacteriol, **174** (20): p. 6444-54. 1992.
273. Hennecke, H., *Nitrogen fixation genes involved in the Bradyrhizobium japonicum-soybean symbiosis*. FEBS Lett, **268** (2): p. 422-6. 1990.
274. Freiberg, C., Fellay, R., Bairoch, A., *et al.*, *Molecular basis of symbiosis between Rhizobium and legumes*. Nature, **387** (6631): p. 394-401. 1997.
275. Cullen, P.J., Bowman, W.C., Hartnett, D.F., *et al.*, *Translational activation by an NtrC enhancer-binding protein*. J Mol Biol, **278** (5): p. 903-14. 1998.
276. Brito, B., Martinez, M., Fernandez, D., *et al.*, *Hydrogenase genes from Rhizobium leguminosarum bv. viciae are controlled by the nitrogen fixation regulatory protein nifA*. Proc Natl Acad Sci U S A, **94** (12): p. 6019-24. 1997.
277. Casalot, L. e Rousset, M., *Maturation of the [NiFe] hydrogenases*. Trends Microbiol, **9** (5): p. 228-37. 2001.
278. Bruckmann, M., Blasco, R., Timmis, K.N., *et al.*, *Detoxification of protoanemonin by dienelactone hydrolase*. J Bacteriol, **180** (2): p. 400-2. 1998.
279. Moiseeva, O.V., Solyanikova, I.P., Kaschabek, S.R., *et al.*, *A new modified ortho cleavage pathway of 3-chlorocatechol degradation by Rhodococcus opacus 1CP: genetic and biochemical evidence*. J Bacteriol, **184** (19): p. 5282-92. 2002.
280. Seibert, V., Thiel, M., Hinner, I.S., *et al.*, *Characterization of a gene cluster encoding the maleylacetate reductase from Ralstonia eutropha 335T, an enzyme recruited for growth with 4-fluorobenzoate*. Microbiology, **150** (Pt 2): p. 463-72. 2004.
281. Holmquist, M., *Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms*. Curr Protein Pept Sci, **1** (2): p. 209-35. 2000.
282. Beveridge, A.J. e Ollis, D.L., *A theoretical study of substrate-induced activation of dienelactone hydrolase*. Protein Eng, **8** (2): p. 135-42. 1995.
283. Park, I.S., Carr, M.B. e Hausinger, R.P., *In vitro activation of urease apoprotein and role of UreD as a chaperone required for nickel metallocenter assembly*. Proc Natl Acad Sci U S A, **91** (8): p. 3233-7. 1994.

284. Toffanin, A., Cadahia, E., Imperial, J., *et al.*, *Characterization of the urease gene cluster from Rhizobium leguminosarum bv. viciae*. Arch Microbiol, **177** (4): p. 290-8. 2002.
285. Sidebotham, R.L., Worku, M.L., Karim, Q.N., *et al.*, *How Helicobacter pylori urease may affect external pH and influence growth and motility in the mucus environment: evidence from in-vitro studies*. Eur J Gastroenterol Hepatol, **15** (4): p. 395-401. 2003.
286. Zhang, W., Culley, D.E., Gritsenko, M.A., *et al.*, *LC-MS/MS based proteomic analysis and functional inference of hypothetical proteins in Desulfovibrio vulgaris*. Biochem Biophys Res Commun, **349** (4): p. 1412-9. 2006.
287. Tanner, S., Shen, Z., Ng, J., *et al.*, *Improving gene annotation using peptide mass spectrometry*. Genome Res. 2007.
288. Choudhary, J.S., Blackstock, W.P., Creasy, D.M., *et al.*, *Matching peptide mass spectra to EST and genomic DNA databases*. Trends Biotechnol, **19** (10 Suppl): p. S17-22. 2001.
289. Choudhary, J.S., Blackstock, W.P., Creasy, D.M., *et al.*, *Interrogating the human genome using uninterpreted mass spectrometry data*. Proteomics, **1** (5): p. 651-67. 2001.

ANEXO



Esquema de um equipamento MALDI-TOF

Esquema de um equipamento ESI-Q-Tof

Tabela do banco proteômico

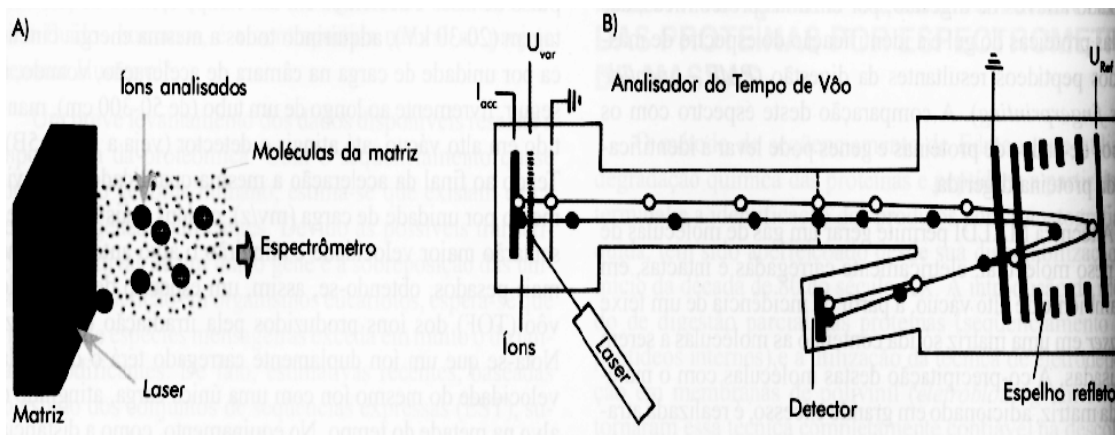


Fig. 1: Esquema do funcionamento de um equipamento MALDI-TOF DE-Reflectron. A) Os íons a serem analisados são extraídos da matriz com um pulso de laser. B) Após breve intervalo de tempo, os íons são acelerados em um campo elétrico de alta voltagem e introduzidos no tubo para análise do tempo de voo; um refletor redireciona a trajetória dos íons na direção oposta, diminuindo a dispersão dos íons com a mesma massa e carga.

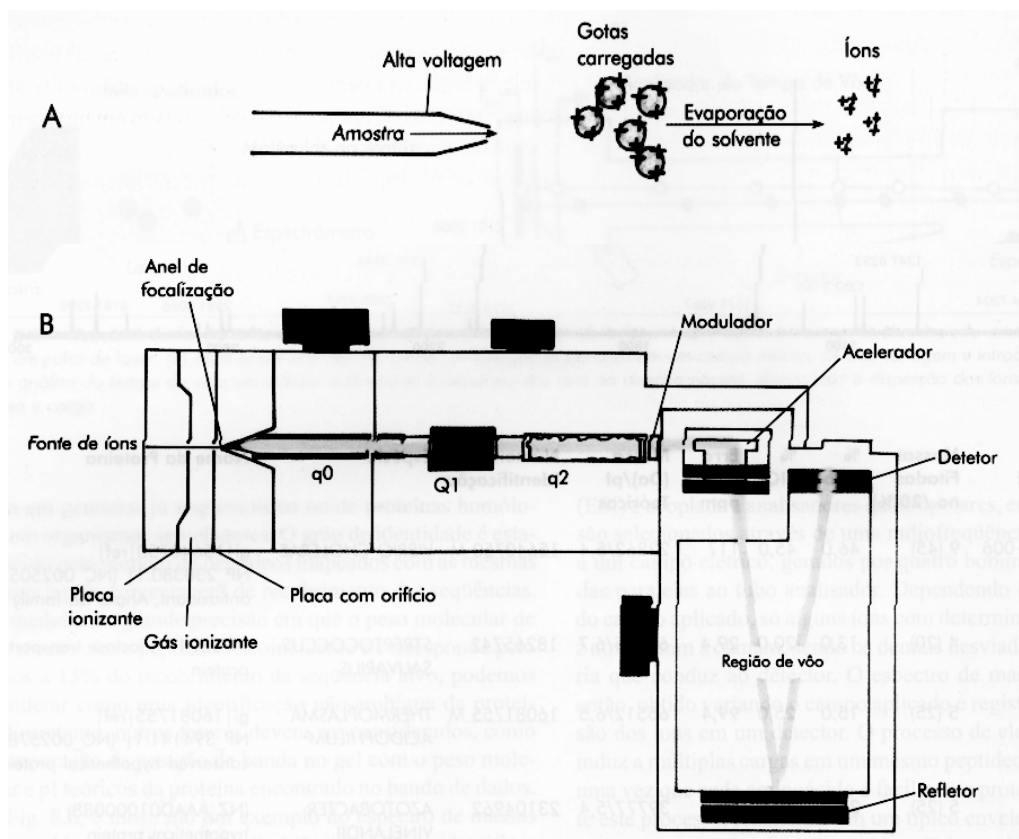


Fig. 2: Esquema de um equipamento ESI-Q-TOF. A) Os íons são produzidos por eletrodispersão em um ambiente atmosférico em presença de um gás ionizante. B) Após o processo de ionização, os íons são conduzidos a um ambiente de alto vácuo através de um orifício, sendo focados no primeiro quadrupolo (q_0). Os peptídeos selecionados em Q_1 são fragmentados em q_2 . Os fragmentos são, então, introduzidos no analisador de tempo de voo através de um modulador pulsado.

TABELA DO BANCO PROTEÔMICO

Tabela 1: Proteínas identificadas no banco proteômico. ORF: código de identificação da ORF no banco do Riogene; 2D-GE: número do *spot* no gel do mapa de referência apresentado na página 55 (fig. 8); Anotação: nome da proteína; Categoria: categoria funcional de acordo com a classificação do *Sanger Institute* para bactérias; EC: *Enzyme Commission number*; Via: via metabólica da qual a proteína participa, as vias estão numeradas de 1 a 40 conforme a tabela 2; Q-Tof: número de peptídeos identificados pela técnica LC/MSMS; MALDI: número de peptídeos identificados por MALDI; Contribuição: contribuição da análise proteômica para a anotação da proteína.

No	2D-GE	ORF	ANOTAÇÃO	CATEGORIA GÊNICA	EC	Via	PEP Q-TOF	PEP MALDI	CONTRIBUIÇÃO
1	c317	2_0271	Bacterioferritin	Adaptação, condições atípicas	-	-	13	3	Confirmação do N-terminal
2	c185	4_00623	putative nucleoside diphosphate	Adaptação, condições atípicas	-	-	-	7	Confirmação de ORF
3	c283, c285	5_0935	Peroxiredoxin	Adaptação, condições atípicas	-	-	-	11	Confirmação de ORF
4		5_1149	Fe-S protein, radical SAM family	Adaptação, condições atípicas	-	-	5	-	Confirmação do N-terminal
5		5_1314	putative Ferritin-like protein	Adaptação, condições atípicas	-	-	2	-	Confirmação de ORF
6	c261	6_0860	Cold shock-like protein cspE	Adaptação, condições atípicas	-	-	11	1	Confirmação de ORF
7	c21	deaD	Cold-shock DEAD box protein A homolog	Adaptação, condições atípicas	-	-	10	8	Confirmação de ORF
8	c21	deaD	Cold-shock DEAD box protein A	Adaptação, condições atípicas	-	-	1	-	Confirmação de ORF
9		djIA	DnaJ-like protein	Adaptação, condições atípicas	-	-	1	-	Confirmação de ORF
10		5_0854	putative lipase protein	Biossíntese de ácidos graxos	-	-	1	-	Confirmação de ORF
11		5_1155	1-deoxy-D-xylulose-5-phosphate synthase	Biossíntese de ácidos graxos	2.2.1.7	35	3	-	Confirmação de frameshift
12	c229	6_01081	glycosyl transferase	Biossíntese de ácidos graxos	-	-	3	3	Confirmação de ORF
13		6_0143	3-oxoacyl-[acyl-carrier-protein] reductase	Biossíntese de ácidos graxos	1.1.1.100	34	4	-	Confirmação de ORF
14	c151	6_0487	IspD/ispF bifunctional enzyme	Biossíntese de ácidos graxos	-	-	1	9	Confirmação de ORF
15		6_0573	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	Biossíntese de ácidos graxos	-	-	3	8	Confirmação de ORF
16	c132	6_0832	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	Biossíntese de ácidos graxos	1.17.4.3	35	1	16	Confirmação do N-terminal

17	6_0855	3-oxoacyl-[acyl-carrier-protein] reductase	Biossíntese de ácidos graxos	1.1.1.100	34	9	-	Confirmação de ORF	
18	6_0876	Geranyltranstransferase	Biossíntese de ácidos graxos	2.5.1.10	35	5	-	Confirmação de ORF	
19	6_0877	1-deoxy-D-xylulose-5-phosphate synthase	Biossíntese de ácidos graxos	2.2.1.7.1	35	4	-	Confirmação de ORF	
20	6_0884	putative esterase	Biossíntese de ácidos graxos	-	-	1	-	Confirmação de ORF	
21	c183	adh	Alcohol dehydrogenase	Biossíntese de ácidos graxos	1.1.1.1	10, 17	8	-	Confirmação de ORF
22	c187	fabD	Malonyl CoA-acyl carrier protein transacylase	Biossíntese de ácidos graxos	2.3.1.39	34	-	11	Confirmação de ORF
23	c71	fabF	3-oxoacyl-[acyl-carrier-protein] synthase 2	Biossíntese de ácidos graxos	-	-	2	3	Confirmação de ORF
24		fabG	3-oxoacyl-[acyl-carrier-protein] reductase	Biossíntese de ácidos graxos	1.1.1.100	34	1	-	Confirmação de ORF
25		fabI	Enoyl-[acyl-carrier-protein] reductase [NADH]	Biossíntese de ácidos graxos	1.3.1.9	34	4	-	Confirmação de ORF
26	c49	2_0029	Dihydroxy-acid dehydratase	Biossíntese de aminoácidos	4.2.1.9	7	14	11	Confirmação de ORF
27		3_0266	putative serine protease	Biossíntese de aminoácidos	-	-	4	-	Confirmação de ORF
28	c158	3_0548	putative cysteine synthase	Biossíntese de aminoácidos	2.5.1.47	12, 13	4	9	Confirmação do N-terminal
29		4_00343	glutamate-cysteine ligase precursor	Biossíntese de aminoácidos	6.3.2.2	1, 6	5	-	Confirmação de ORF
30		4_00552	putative shikimate kinase	Biossíntese de aminoácidos	2.7.1.71	4	2	-	Confirmação de ORF
31		5_0472	Histidinol-phosphate aminotransferase	Biossíntese de aminoácidos	2.6.1.9	4, 10, 14	1	-	Confirmação de ORF
32		5_1649	Homoserine dehydrogenase	Biossíntese de aminoácidos	1.1.1.3	2, 8	4	-	Confirmação de ORF
33		6_0287	Succinate-semialdehyde dehydrogenase [NADP+]	Biossíntese de aminoácidos	1.2.1.16	1, 10,	2	-	Confirmação de ORF

34		6_0288	putative branched-chain-amino-acid aminotransferase	Biossíntese de aminoácidos	2.6.1.42	22 7	6	-	Confirmação de ORF
35	c136	6_0294	Alcohol dehydrogenase	Biossíntese de aminoácidos	1.1.1.1	10, 17	23	9	Confirmação de ORF
36		6_0311	L-amino-acid oxidase precursor	Biossíntese de aminoácidos	1.4.3.2	3, 4, 5, 10, 14, 15	1	-	Confirmação de ORF
37		6_0404	Catalase HPII	Biossíntese de aminoácidos	-	-	1	-	Confirmação de ORF
38	c44, c61	6_0422	2-isopropylmalate synthase	Biossíntese de aminoácidos	2.3.3.13	7, 19, 20	3	17	Confirmação de ORF
39		6_0440	Spermidine synthase	Biossíntese de aminoácidos	2.5.1.16	3, 9, 11, 16	4	-	Confirmação de ORF
40	c149	6_0442	glycine cleavage system aminomethyltransferase T	Biossíntese de aminoácidos	-	-	6	4	Confirmação do C-terminal
41	c204	6_0620	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	Biossíntese de aminoácidos	2.3.1.117	8	4	6	Confirmação do C-terminal
42		6_0640	Glutamate synthase [NADPH] small chain	Biossíntese de aminoácidos	1.4.1.13	1, 30	1	-	Confirmação de ORF
43		6_0649	Glutathione synthetase	Biossíntese de aminoácidos	6.3.2.3	1, 6	3	-	Confirmação de ORF
44	c79	6_0663	Threonine synthase	Biossíntese de aminoácidos	4.2.3.1	2	-	7	Confirmação de ORF
45		6_0741	Omega-amino acid--pyruvate aminotransferase	Biossíntese de aminoácidos	-	-	2	-	Confirmação de ORF
46	c250	6_0777	3-mercaptopyruvate sulfurtransferase	Biossíntese de aminoácidos	2.8.1.2	13	1	11	Confirmação de ORF
47	c86, c87	6_0817	Enolase	Biossíntese de aminoácidos	4.2.1.11	4, 17,	9	9	Confirmação de ORF

48	c96	6_0831	Histidyl-tRNA synthetase	Biossíntese de aminoácidos	6.1.1.21	30 39	1	11	Confirmação de ORF
49		6_0835	Aspartokinase	Biossíntese de aminoácidos	2.7.2.4	2, 8	1	-	Confirmação de ORF
50	c249	6_0862	Dihydrodipicolinate synthase	Biossíntese de aminoácidos	4.2.1.52	8	-	7	Confirmação de ORF
51		6_0867	D-3-phosphoglycerate dehydrogenase	Biossíntese de aminoácidos	1.1.1.95	2	1	-	Confirmação de ORF
52	c67	6_0880	Argininosuccinate synthase	Biossíntese de aminoácidos	6.3.4.5	5, 9, 11	5	10	Confirmação de ORF
53	c19	6_0932	PrkA serine protein kinase	Biossíntese de aminoácidos	-	-	31	13	Confirmação do N-terminal
54		6_0935	Methionyl-tRNA synthetase	Biossíntese de aminoácidos	6.1.1.10	3, 12, 39	2	-	Confirmação de ORF
55		6_0971	Peptide methionine sulfoxide reductase msrA	Biossíntese de aminoácidos	1.8.4.11	-	1	-	Confirmação de ORF
56		ahcY	putative Adenosylhomocysteinase	Biossíntese de aminoácidos	3.3.1.1	3, 12	1	-	Confirmação de ORF
57		aldA	Aldehyde dehydrogenase	Biossíntese de aminoácidos	-	-	3	-	Confirmação de ORF
58		argJ	Arginine biosynthesis bifunctional protein argJ	Biossíntese de aminoácidos	-	-	9	-	Confirmação de ORF
59	c134	asd	putative aspartate-semialdehyde dehydrogenase	Biossíntese de aminoácidos	1.2.1.11	2, 8	12	1	Confirmação do N-terminal
60		cysB	Cysteine synthase	Biossíntese de aminoácidos	2.5.1.47	-	4	-	Confirmação de ORF
61		dapB	putative Dihydrodipicolinate reductase	Biossíntese de aminoácidos	1.3.1.26	8, 16, 38	3	1	Confirmação de ORF
62		gabD	Succinate-semialdehyde dehydrogenase	Biossíntese de aminoácidos	1.2.1.24	1, 22	2	-	Confirmação de ORF
63	c47, c48	glnA	Glutamine synthetase	Biossíntese de aminoácidos	6.3.1.2	1, 30,	53	14	Confirmação de ORF

64		glyA	Serine hydroxymethyltransferase	Biossíntese de aminoácidos	2.1.2.1	40, 2, 29, 32	5	-	Confirmação de ORF
65		hisA	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino imidazole-4-carboxamide isomerase	Biossíntese de aminoácidos	-	-	2	-	Confirmação de ORF
66		ilvC	Ketol-acid reductoisomerase	Biossíntese de aminoácidos	1.1.1.86	7	5	-	Confirmação de ORF
67	c51, c53, c54, c55, c56	katA	Catalase	Biossíntese de aminoácidos	1.1.1.6	15, 29	29	12	Confirmação de ORF
68		kbl	2-amino-3-ketobutyrate coenzyme A ligase	Biossíntese de aminoácidos	2.3.1.29	2	1	-	Confirmação de ORF
69		lepA	GTP-binding protein lepA	Biossíntese de aminoácidos	-	-	1	-	Confirmação de ORF
70		leuC	3-isopropylmalate dehydratase large subunit	Biossíntese de aminoácidos	4.2.1.33	7	3	-	Confirmação de ORF
71		leuD	3-isopropylmalate dehydratase small subunit	Biossíntese de aminoácidos	4.2.1.33	7	1	-	Confirmação de ORF
72		metW	putative methionine biosynthesis MetW	Biossíntese de aminoácidos	-	-	1	-	Confirmação de ORF
73		proA	Gamma-glutamyl phosphate reductase	Biossíntese de aminoácidos	1.2.1.38	11	1	-	Confirmação de ORF
74		tkrA	2-ketogluconate reductase	Biossíntese de aminoácidos	1.1.1.215	-	1	-	Confirmação de ORF
75	c213	trpA	putative Tryptophan synthase alpha chain	Biossíntese de aminoácidos	4.2.1.20	4, 20, 31, 40	2	7	Confirmação de ORF
76		trpB	putative Tryptophan synthase beta chain	Biossíntese de aminoácidos	4.2.1.20	4, 20, 31,	2	-	Confirmação de ORF

77	4_00256	putative inositol monophosphatase	Biossíntese de cofatores	3.1.3.25	40 36	1	-	Confirmação de ORF
78	6_0197	3-alpha-(or 20-beta)-hydroxysteroid dehydrogenase	Biossíntese de cofatores	-	-	4	-	Confirmação de ORF
79	6_0206	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferas	Biossíntese de cofatores	2.1.1.14	3	52	-	Confirmação de ORF
80	6_0245	Adenine deaminase	Biossíntese de cofatores	3.5.4.2	37	3	-	Confirmação de ORF
81	6_0248	Arginine deiminase	Biossíntese de cofatores	3.5.3.6	9	1	-	Confirmação de ORF
82	6_0412	2-octaprenyl-6-methoxyphenol hydroxylase	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
83	6_0462	Gamma-glutamyltranspeptidase precursor	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
84	6_0476	Glutathione S-transferase	Biossíntese de cofatores	2.5.1.18	6	1	-	Confirmação de ORF
85	c85 6_0539	Glutathione reductase	Biossíntese de cofatores	1.8.1.7	1, 6	-	12	Confirmação do N-terminal
86	6_0541	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (DAHP synthetase, phenylalanine repressible) [EC:2.5.1.54]	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
87	6_0574	Probable molybdopterin binding	Biossíntese de cofatores	-	-	2	-	Confirmação de ORF
88	6_0806	Molybdenum cofactor biosynthesis protein C	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
89	6_0836	3-demethylubiquinone-9 3-methyltransferase	Biossíntese de cofatores	2.1.1.64	-	3	-	Confirmação de ORF
90	6_0866	Inositol-1-monophosphatase	Biossíntese de cofatores	3.1.3.25	36	1	-	Confirmação de ORF

91		6_0953	Thiamine-phosphate pyrophosphorylase	Biossíntese de cofatores	2.5.1.3	-	3	-	Confirmação de ORF
92		6_0954	Thiazole biosynthesis protein thiG	Biossíntese de cofatores	-	-	5	-	Confirmação de ORF
93	c165r	6_0956	thiamine biosynthesis oxidoreductase thiO	Biossíntese de cofatores	-	-	7	4	Confirmação de ORF
94		6_0967	Ubiquinone/menaquinone biosynthesis methyltransferase ubi	Biossíntese de cofatores	-	-	7	-	Confirmação de ORF
95		accC	putative Biotin carboxylase protein	Biossíntese de cofatores	6.3.4.14	34, 37	4	-	Confirmação de ORF
96		apbA	Putative 2-dehydropantoate 2-reductase	Biossíntese de cofatores	1.1.1.169	-	1	-	Confirmação de ORF
97		cobC	Alpha-ribazole phosphatase	Biossíntese de cofatores	3.1.3.73	33	1	-	Confirmação de ORF
98		folE	GTP cyclohydrolase I	Biossíntese de cofatores	3.5.4.16	-	1	-	Confirmação de ORF
99		hemE	Uroporphyrinogen decarboxylase	Biossíntese de cofatores	4.1.1.37	33	3	-	Confirmação de ORF
100		hemH	putative ferrochelataze	Biossíntese de cofatores	4.99.1.1	33	2	-	Confirmação de ORF
101	c302	moaB	Molybdenum cofactor biosynthesis protein B	Biossíntese de cofatores	-	-	2	8	Confirmação de ORF
102		moeB	putative molybdopterin biosynthesis protein moeB	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
103		panB	putative 3-methyl-2-oxobutanoate hydroxymethyltransferase	Biossíntese de cofatores	2.1.2.11	-	1	-	Confirmação de ORF
104		pdxJ	putative pyridoxal phosphate biosynthetic protein	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
105		pncA	putative pyrazinamidase/nicotinamidase	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF
106	c194	pqqB	Coenzyme PQQ synthesis protein B	Biossíntese de cofatores	-	-	1	5	Confirmação de ORF

107	pqqC	putative Pyrroloquinoline-quinone synthase	Biossíntese de cofatores	1.3.3.11	-	2	-	Confirmação de ORF	
108	ribD	putative riboflavin biosynthesis protein	Biossíntese de cofatores	2.5.1.9	-	1	-	Confirmação de ORF	
109	c294	ribE	putative riboflavin synthase alpha chain	Biossíntese de cofatores	2.5.1.9	-	3	3	Confirmação de ORF
110	suhB	Inositol-1-monophosphatase	Biossíntese de cofatores	3.1.3.25	36	4	-	Confirmação de ORF	
111	thiC	Thiamine biosynthesis protein	Biossíntese de cofatores	-	-	1	-	Confirmação de ORF	
112	trxB	Thioredoxin reductase	Biossíntese de cofatores	1.8.1.9	38	1	-	Confirmação de ORF	
113	c279	wrbA	putative flavoprotein	Biossíntese de cofatores	-	-	-	6	Confirmação de ORF
114	6_01001	Acetoin(diacetyl) reductase	Biossíntese de nucleotídeos	1.1.1.5	-	1	-	Confirmação de ORF	
115	c320	6_0733	Nucleoside diphosphate kinase	Biossíntese de nucleotídeos	2.7.4.6	37, 38	4	2	Confirmação de ORF
116	6_0913	Ribonucleoside-diphosphate reductase large subunit	Biossíntese de nucleotídeos	1.17.4.1	37, 38	1	-	Confirmação de ORF	
117	3_0035	putative pyridine nucleotide-disulphide oxidoreductase	Biossíntese de pirimidinas	1.18.1.3	-	5	-	Confirmação de ORF	
118	3_0586	Dihydropyrimidine dehydrogenase	Biossíntese de pirimidinas	1.3.1.2	-	1	-	Confirmação de ORF	
119	6_0814	CTP synthase	Biossíntese de pirimidinas	6.3.4.2	38	4	-	Confirmação de ORF	
120	c135	nadA	Quinolinate synthetase A	Biossíntese de pirimidinas	-	-	1	5	Confirmação de ORF
121	nadC	Nicotinate-nucleotide pyrophosphorylase	Biossíntese de pirimidinas	2.4.2.19	-	4	-	Confirmação de ORF	
122	thyX	Thymidylate synthase thyX	Biossíntese de pirimidinas	2.1.1.148	32, 38	1	-	Confirmação de ORF	
123	c138	trxB	Thioredoxin reductase	Biossíntese de pirimidinas	1.8.1.9	-	-	5	Confirmação de ORF
124	c138	trxB	Thioredoxin reductase	Biossíntese de	-	32,	-	5	Confirmação de ORF

				pirimidinas		38			
125	c101	6_0501	Adenylosuccinate synthetase	Biossíntese de purinas	6.3.4.4	5, 37	4	11	Confirmação de ORF
126		6_0694	5'-methylthioadenosine phosphorylase	Biossíntese de purinas	2.4.2.28	3	2	-	Confirmação de ORF
127		6_0785	Phosphoribosylformylglycine midine synthase II	Biossíntese de purinas	6.3.5.3	37	2	-	Confirmação do C-terminal
128	c263	6_0786	Phosphoribosylformylglycine midine synthase I	Biossíntese de purinas	6.3.5.3	37	3	-	Confirmação de ORF
129		6_0789	Phosphoribosylaminoimidazole-succinocarboxamide synthase (Biossíntese de purinas	6.3.2.6	37	1	-	Confirmação de ORF
130		6_0908	Orotate phosphoribosyltransferase	Biossíntese de purinas	2.4.2.10	38	1	-	Confirmação de ORF
131	c57	6_0924	Inosine-5'-monophosphate dehydrogenase	Biossíntese de purinas	1.1.1.205	37	2	10	Confirmação de ORF
132		gpt	Xanthine phosphoribosyltransferase	Biossíntese de purinas	2.4.2.22	37	1	-	Confirmação de ORF
133		guaA	GMP synthase	Biossíntese de purinas	6.3.4.1	-	7	-	Confirmação de ORF
134		prs	Ribose-phosphate pyrophosphokinase	Biossíntese de purinas	2.7.6.1	18, 37	2	-	Confirmação de ORF
135		purH	putative bifunctional purine biosynthesis protein purH	Biossíntese de purinas	3.5.4.10	37	3	-	Confirmação de ORF
136	c274	6_0959	Phosphoglycolate phosphatase	Ciclo do glioxilato	3.1.3.18	27	2	6	Confirmação de ORF
137		6_0756	Isocitrate dehydrogenase	Ciclo TCA	1.1.1.41	20	-	8	Confirmação de ORF
138		6_0906	Citrate synthase	Ciclo TCA	2.3.3.1	7, 19, 20, 27	2	-	Confirmação de ORF
139	c50	fumA	Fumarate hydratase class	Ciclo TCA	4.2.1.2	-	21	-	Confirmação de ORF
140		icd	Isocitrate dehydrogenase [NADP]	Ciclo TCA	1.1.1.42	6, 20, 31	3	-	Confirmação de ORF
141		4_00019	putative Polyphosphate kinase	Cinases	2.7.4.1	-	1	-	Confirmação de ORF
142		5_1124	Inosine-guanosine kinase	Cinases	2.7.1.73	37	1	-	Confirmação de ORF
143		6_0148	Ribulokinase	Cinases	2.7.1.16	23	6	-	Confirmação de ORF

144		6_0456	Phosphoglycerate kinase	Cinases	2.7.2.3	17, 28	4	-	Confirmação de ORF
145		6_0603	Uridylate kinase	Cinases	2.7.4.22	-	11	-	Confirmação de ORF
146	c179	rbsK	Ribokinase	Cinases	2.7.1.15	18	1	1	Confirmação de ORF
147		4_00220	Uracil phosphoribosyltransferase	Conversão e biossíntese nucleotídeos-açúcares	2.4.2.9	38	2	-	Confirmação de ORF
148		6_0382	Ribulose-phosphate 3-epimerase	Conversão e biossíntese nucleotídeos-açúcares	5.1.3.1	18, 23, 28	1	-	Confirmação de ORF
149	c267	cmk	Cytidylate kinase	Conversão e biossíntese nucleotídeos-açúcares	2.7.4.14	38	2	7	Confirmação do N-terminal
150		galE	UDP-glucose 4-epimerase	Conversão e biossíntese nucleotídeos-açúcares	5.1.3.2	24, 26	1	-	Confirmação de ORF
151		rmlA	Glucose-1-phosphate thymidyltransferase 1	Conversão e biossíntese nucleotídeos-açúcares	2.7.7.24	26, 36	3	-	Confirmação de ORF
152		rmlA	Glucose-1-phosphate thymidyltransferase	Conversão e biossíntese nucleotídeos-açúcares	2.7.7.24	26, 36	3	-	Confirmação de ORF
153	c96	6_0742	N-carbamoyl-L-amino acid amidohydrolase	Degradação de macromoléculas	-	-	-	5	Confirmação de ORF
154	c260	4_00338	Putative carboxymethylenebutenolidase	Degradação de pequenas moléculas	-	-	-	5	Confirmação de ORF
155	c305	5_0933	putative carboxymuconolactone decarboxylase	Degradação de pequenas moléculas	4.1.1.44	-	3	4	Confirmação do C-terminal
156		6_0578	Vanillate O-demethylase oxygenase subunit	Degradação de pequenas moléculas	-	-	1	-	Confirmação de ORF
157		6_0695	putative dienelactone hydrolase	Degradação de pequenas moléculas	3.1.1.45	-	1	-	Confirmação de ORF
158		5_1399	putative lipase protein	Degradação de polisacarídeos	-	-	1	-	Confirmação de ORF

159	c63	3_0305	putative glutamate carboxypeptidase protein	Degradação de proteínas, peptídeos, glicoproteínas	3.4.17.11	-	-	16	Confirmação do C-terminal
160		3_0414	putative glycoprotease protein	Degradação de proteínas, peptídeos, glicoproteínas	3.4.24.57	-	1	-	Confirmação de ORF
161	c156	3_0465	putative dipeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	1	7	Confirmação de ORF
162		3_0524	putative aminopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	3	-	Confirmação de ORF
163		4_00064	putative insulinase	Degradação de proteínas, peptídeos, glicoproteínas	3.4.24.56	-	6	18	Confirmação do C-terminal
164	c76	4_00213	putative carboxy-terminal protease protein	Degradação de proteínas, peptídeos, glicoproteínas	-	-	5	12	Confirmação de ORF
165	c89	4_00258	putative cytosol aminopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	1	6	Confirmação de ORF
166		4_00503	putative dipeptidyl-peptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	16	4	Confirmação de ORF
167		4_00565	putative oligopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	1	-	Confirmação de ORF
168		4_00631	putative serine carboxypeptidase	Degradação de proteínas, peptídeos, glicoproteínas	3.4.16.5	-	4	7	Confirmação de ORF
169		4_00800	putative metallopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	2	-	Confirmação de ORF
170	c245	5_1230	putative protease	Degradação de proteínas, peptídeos,	-	-	1	3	Confirmação de ORF

171	c146	5_1309	Dipeptidyl peptidase	glicoproteínas Degradação de proteínas, peptídeos, glicoproteínas	-	-	-	10	Confirmação do C-terminal
172		6_0295	Aminobenzoyl-glutamate utilization protein	Degradação de proteínas, peptídeos, glicoproteínas	-	-	2	-	Confirmação de ORF
173		6_0331	glycyl aminopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	14	-	Confirmação de ORF
174	c93	6_0385	Aminopeptidase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	-	8	Confirmação de ORF
175		6_0642	Group B oligopeptidase pepB	Degradação de proteínas, peptídeos, glicoproteínas	-	-	16	-	Confirmação de ORF
176	c66	6_0661	peptidase U62, modulator of DNA gyrase	Degradação de proteínas, peptídeos, glicoproteínas	-	-	2	9	Confirmação de ORF
177		6_0662	peptidase, family M16	Degradação de proteínas, peptídeos, glicoproteínas	-	-	2	-	Confirmação de ORF
178		6_0666	Thermostable carboxypeptidase 1	Degradação de proteínas, peptídeos, glicoproteínas	-	-	1	-	Confirmação de ORF
179		6_0743	D-hydantoinase (Dihydropyrimidinase) (DHPase)	Degradação de proteínas, peptídeos, glicoproteínas	3.5.2.2	16, 38	1	-	Confirmação de ORF
180		6_0757	Peptidase M24	Degradação de proteínas, peptídeos, glicoproteínas	-	-	6	-	Confirmação de ORF
181		clpA	ATP-dependent Clp protease ATP-binding subunit clpA	Degradação de proteínas, peptídeos, glicoproteínas	3.4.21.92	-	9	-	Confirmação de ORF
182	c242a	clpP	ATP-dependent Clp protease	Degradação de	3.4.21.92	-	4	1	Confirmação de ORF

			proteolytic subunit 1	proteínas, peptídeos, glicoproteínas					
183	c224	clpP	ATP-dependent Clp protease proteolytic subunit 2	Degradação de proteínas, peptídeos, glicoproteínas	3.4.21.92	-	7	9	Confirmação de ORF
184		clpX	ATP-dependent Clp protease ATP-binding subunit	Degradação de proteínas, peptídeos, glicoproteínas	3.4.21.92	-	3	-	Confirmação de ORF
185		hslV	ATP-dependent protease hslV	Degradação de proteínas, peptídeos, glicoproteínas	-	-	1	-	Confirmação de ORF
186		1_011	short-chain dehydrogenase	Desconhecida	-	-	-	10	Confirmação de frameshift
187		2_0316	hypothetical protein	Desconhecida	-	-	14	-	Confirmação de ORF
188		2_0355	hypothetical protein	Desconhecida	-	-	1	-	Confirmação de ORF
189		3_0224	hypothetical protein	Desconhecida	-	-	2	-	Confirmação de ORF
190		3_0269	hypothetical protein	Desconhecida	-	-	4	-	Confirmação de ORF
191	c192	3_0376	hypothetical protein	Desconhecida	-	-	3	5	Confirmação de ORF
192		3_0385	putative Aldo/keto reductase family	Desconhecida	-	-	1	-	Confirmação de ORF
193		3_0502	hypothetical protein	Desconhecida	-	-	2	-	Confirmação de ORF
194		3_0517	hypothetical protein	Desconhecida	-	-	1	-	Confirmação de ORF
195		3_0690	hypothetical protein	Desconhecida	-	-	3	-	Confirmação de ORF
196		4_00091	hypothetical protein	Desconhecida	-	-	2	-	Confirmação de ORF
197	c251	4_00264	putative oxidoreductase	Desconhecida	-	-	1	7	Confirmação de ORF
198		4_00315	putative D-mannonate oxidoreductase	Desconhecida	-	-	2	-	Confirmação de ORF
199		4_00322	putative dehydrogenase	Desconhecida	-	-	1	-	Confirmação de ORF
200		4_00522	hypothetical protein	Desconhecida	-	-	5	-	Confirmação de ORF
201		4_00567	putative oxidoreductase protein	Desconhecida	-	-	1	-	Confirmação de ORF
202		4_00748	hypothetical protein	Desconhecida	-	-	1	-	Confirmação de ORF
203		4_368ne w369	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita ?

204		4_3new 4	hypothetical protein	Desconhecida	-	-	2	-	Assinalamento de ORF não-predita
205		4_419ne w420	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita ?
206		4_454ne w455	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita ?
207		4_921ne w925	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita?
208	c105	5_1371	putative Major royal jelly-like protein	Desconhecida	-	-	-	10	Confirmação de ORF
209		5_1371	putative Major royal jelly-like protein	Desconhecida	-	-	-	10	Confirmação de ORF
210	c269	5_1668	hypothetical protein	Desconhecida	-	-	-	3	Confirmação de ORF
211		5_215ne w216	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita??
212		5_new2	hypothetical protein	Desconhecida	-	-	2	-	Assinalamento de ORF não-predita?
213		6_0208	PRC-barrel	Desconhecida	-	-	1	-	Confirmação de ORF
214		6_0564	oxidoreductase	Desconhecida	-	-	1	-	Confirmação de ORF
215		6_0618	Probable GTP-binding protein engB	Desconhecida	-	-	1	-	Confirmação de ORF
216		6_0664	hypothetical protein	Desconhecida	-	-	4	-	Confirmação de ORF
217		6_0664	hypothetical protein	Desconhecida	-	-	-	3	Confirmação de ORF
218		6_0800	putative cyclase/dehydrase protein	Desconhecida	-	-	1	-	Confirmação de ORF
219		6_0899	putative thioredoxin	Desconhecida	-	-	2	-	Confirmação de ORF
220	c206	6_430ne w431	hypothetical protein	Desconhecida	-	-	-	4	Assinalamento de ORF não-predita
221		6_832ne w833	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita
222		6_975ne w976	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de ORF não-predita
223		6_dnajn ewcobt	hypothetical protein	Desconhecida	-	-	2	-	Assinalamento de ORF não-predita
224		6_new	hypothetical protein	Desconhecida	-	-	1	-	Assinalamento de

225	new	hypothetical protein	Desconhecida	-	-	2	-	ORF não-predita?? Assinalamento de ORF não-predita?
226	tldD	putative modulator of DNA gyrase protein	Desconhecida	-	-	3	-	Confirmação de ORF
227	xxxx	hypothetical protein	Desconhecida	-	-	3	-	Confirmação de ORF
228	c160	6_0455 Cell division protein ftsY	Divisão celular	-	-	-	7	Confirmação de ORF
229	c256	6_0474 Septum site-determining protein minD	Divisão celular	-	-	5	12	Confirmação de ORF
230	c160	ftsY Cell division protein ftsY	Divisão celular	-	-	-	7	Confirmação de ORF
231		ftsZ Cell division protein ftsZ	Divisão celular	-	-	18	-	Confirmação do C- terminal
232	maf	putative inhibitor of septum formation, Maf-like protein	Divisão celular	-	-	1	-	Confirmação de ORF
233	c119	2_0025 putative secretion protein HlyD	Envelope celular	-	-	3	6	Confirmação de ORF
234		4_00108 putative outer membrane efflux protein	Envelope celular	-	-	1	-	Confirmação de ORF
235		4_00183 putative outer membrane protein	Envelope celular	-	-	1	-	Confirmação de ORF
236		4_00610 putative outer membrane protein	Envelope celular	-	-	1	-	Confirmação de ORF
237		5_0081 putative surface antigen	Envelope celular	-	-	2	-	Confirmação do C- terminal
238	c215	5_1023 putative membrane protein	Envelope celular	-	-	-	9	Confirmação de ORF
239		5_1236 putative surface antigen	Envelope celular	-	-	1	-	Confirmação de ORF
240	c120	6_0329 secretion protein HlyD	Envelope celular	-	-	-	11	Confirmação de ORF
241		6_0329 secretion protein HlyD	Envelope celular	-	-	-	11	Confirmação de ORF
242		6_0351 Putative polysaccharide export protein	Envelope celular	-	-	1	-	Confirmação de ORF
243		6_0355 Capsule polysaccharide export ATP-binding protein ctrD	Envelope celular	-	-	1	-	Confirmação de ORF
244	c109, c152, c203,	6_0569 Outer membrane lipoprotein	Envelope celular	-	-	51	10	Confirmação do C- terminal

245	c204, c25	6_0609	Outer membrane protein assembly factor yaeT precursor	Envelope celular	-	-	7	-	Confirmação de ORF
246		6_0610	Outer membrane protein	Envelope celular	-	-	4	-	Confirmação de ORF
247	c308	6_0681	Protein-export protein secB	Envelope celular	-	-	5	1	Confirmação de ORF
248		czcB	Cation efflux system protein	Envelope celular	-	-	-	11	Confirmação de ORF
249		imp	Organic solvent tolerance protein precursor	Envelope celular	-	-	1	-	Confirmação de ORF
250	c193	murB	UDP-N-acetylenolpyruvoylglucosamin e reductase	Envelope celular	1.1.1.158	-	5	-	Confirmação de ORF
251	c313a	omp	Outer membrane lipoprotein omp16	Envelope celular	-	-	5	-	Confirmação de ORF
252		OmpW	putative Outer membrane protein	Envelope celular	-	-	32	-	Confirmação de ORF
253		oprJ	putative outer membrane protein	Envelope celular	-	-	5	-	Confirmação de ORF
254	c235, c236	pbp	putative phosphate-binding periplasmic protein precursor	Envelope celular	-	-	-	8	Confirmação de ORF
255		rbsB	D-ribose-binding periplasmic protein precursor	Envelope celular	-	-	-	6	Confirmação de ORF
256		4_00275	Phosphatidylserine decarboxylase proenzyme	Fosfolipídeos	4.1.1.65	2	1	-	Confirmação de ORF
257	c144, c168, c169, c170	6_0457	Glyceraldehyde-3-phosphate dehydrogenase	Glicólise	1.2.1.13	28	31	9	Confirmação de ORF
258		6_0873	Glucokinase	Glicólise	2.7.1.2	17, 24, 25, 36	3	-	Confirmação de ORF
259		6_0975	Glucokinase	Glicólise	2.7.1.2	17, 24,	3	-	Confirmação de ORF

260	c131	galM	Aldose 1-epimerase precursor	Glicólise	5.1.3.3	25, 36 17	-	7	Confirmação de ORF
261	c163	glpX	Fructose-1,6-bisphosphatase class II	Glicólise	3.1.3.11	17, 18, 20, 28	-	10	Confirmação de ORF
262	c54	gpmB	Probable phosphoglycerate mutase gpmB	Glicólise	5.4.2.1	17	1	-	Confirmação de ORF
263		podA	putative pyruvate, phosphate dikinase	Glicólise	2.7.9.1	19, 28	2	-	Confirmação de ORF
264	c46	pykA	putative pyruvate kinase	Glicólise	2.7.1.40	17, 19, 28, 37	23	6	Confirmação do N-terminal
265		4_00217	2,3-bisphosphoglycerate-independent phosphoglycerate mutas	Gliconeogênese	5.4.2.4	17	17	9	Confirmação de ORF
266	c39, c60	pgm	Phosphoglucomutase	Gliconeogênese	5.4.2.2	17, 18, 24, 25, 36	7	15	Confirmação de N e C-terminal
267	c40, c59, c61	zwf	putative glucose-6-phosphate 1-dehydrogenase	Gliconeogênese	1.1.1.49	6, 18	34	2	Confirmação de ORF
268		2_0049	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
269		2_0065	conserved hypothetical protein	Hipotética conservada	-	-	4	-	Confirmação de ORF
270		3_0066	conserved hypothetical protein	Hipotética conservada	-	-	2	-	Confirmação de ORF
271		3_0109	conserved hypothetical protein	Hipotética conservada	-	-	4	-	Confirmação de ORF
272	c212	3_0362	conserved hypothetical protein	Hipotética conservada	-	-	1	9	Confirmação de ORF
273	c175	3_0408	conserved hypothetical protein	Hipotética conservada	-	-	-	6	Confirmação de ORF
274	c175	3_0408	conserved hypothetical protein	Hipotética conservada	-	-	-	8	Confirmação de ORF
275		4_00085	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
276		4_00176	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF

277		4_00202	conserved hypothetical protein	Hipotética conservada	-	-	-	7	Confirmação de ORF
278		4_00331	conserved hypothetical protein	Hipotética conservada	-	-	3	-	Confirmação de ORF
279	c323	4_00359	conserved hypothetical protein	Hipotética conservada	-	-	-	6	Confirmação de ORF
280		4_00570	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
281	c217	4_00613	conserved hypothetical protein	Hipotética conservada	-	-	-	3	Confirmação de ORF
282	c217	4_00613	conserved hypothetical protein	Hipotética conservada	-	-	-	4	Confirmação de ORF
283		4_00648	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
284		4_00713	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
285		4_00792	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
286		5_0107	conserved hypothetical protein	Hipotética conservada	-	-	7	-	Confirmação do C-terminal
287		5_0110	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
288		5_0280	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
289		5_0299	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
290	c311	5_0473	conserved hypothetical protein	Hipotética conservada	-	-	2	7	Confirmação de ORF
291		5_0672	conserved hypothetical protein	Hipotética conservada	-	-	2	-	Confirmação de ORF
292		5_0824	conserved hypothetical protein	Hipotética conservada	-	-	2	-	Confirmação de ORF
293		5_1122	conserved hypothetical protein	Hipotética conservada	-	-	20	-	Confirmação de ORF
294		5_1180	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
295		5_1248	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
296	c214	5_1270	conserved hypothetical protein	Hipotética conservada	-	-	1	5	Confirmação do C-terminal
297		5_1396	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
298	c303	5_1439	conserved hypothetical protein	Hipotética conservada	-	-	-	5	Confirmação de ORF
299	c303	5_1439	conserved hypothetical protein	Hipotética conservada	-	-	-	6	Confirmação de ORF
300	c209	5_1457	conserved hypothetical protein	Hipotética conservada	-	-	3	4	Confirmação de ORF
301	c336, c337	5_1462	conserved hypothetical protein	Hipotética conservada	-	-	-	7	Confirmação de N e C-terminal
302	c336, c337	5_1462	conserved hypothetical protein	Hipotética conservada	-	-	-	6	Confirmação de ORF
303		6_01006	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
304		6_01061	conserved hypothetical protein	Hipotética conservada	-	-	3	-	Confirmação de ORF
305		6_0484	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF

306		6_0517	conserved hypothetical protein	Hipotética conservada	-	-	4	-	Confirmação de ORF
307		6_0542	conserved hypothetical protein	Hipotética conservada	-	-	2	-	Confirmação de ORF
308	c167r	6_0548	conserved hypothetical protein	Hipotética conservada	-	-	-	11	Confirmação de ORF
309	c167r	6_0548	conserved hypothetical protein	Hipotética conservada	-	-	-	11	Confirmação de ORF
310		6_0579	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
311		6_0790	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
312	c181	6_0859	conserved hypothetical protein	Hipotética conservada	-	-	7	5	Confirmação do C-terminal
313		6_0921	conserved hypothetical protein	Hipotética conservada	-	-	1	-	Confirmação de ORF
314		6_0845	Lipoyltransferase	Lipoate	2.3.1.181	-	4	-	Confirmação de ORF
315		6_0613	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferas	Lipopolissacarídeos	2.3.1.129	-	1	-	Confirmação de ORF
316	c246	rbcC	dTDP-4-dehydrorhamnose 3,5-epimerase	Lipopolissacarídeos	5.1.3.13	26, 36	1	8	Confirmação de ORF
317	c279	wrbA	putative flavoprotein	Metabolismo central e intermediário	-	-	-	6	Confirmação de ORF
318	c226	6_0566	Carbonic anhydrase 1	Metabolismo de energia	4.2.1.1	-	7	3	Confirmação de ORF
319		adk	Adenylate kinase	Metabolismo de energia	2.7.4.11	37	9	-	Confirmação de ORF
320		atpC	ATP synthase epsilon chain	Metabolismo de energia	-	-	-	6	Confirmação de ORF
321		galM	Aldose 1-epimerase precursor	Metabolismo de energia	-	-	-	10	Confirmação de ORF
322	c143	mauB	Methylamine dehydrogenase heavy chain precursor	Metabolismo de energia	1.4.99.3	29	-	8	Confirmação do N-terminal
323	c59	4_00517	Phosphomannomutase	Metabolismo de glicose	5.4.2.8	21	-	13	Confirmação do N-terminal
324		6_0158	Acetoin(diacetyl) reductase	Metabolismo de glicose	1.1.1.5	22	2	-	Confirmação do N-terminal
325	c197	6_0844	UTP--glucose-1-phosphate uridylyltransferase	Metabolismo de glicose	2.7.7.9	23, 24, 25, 26	1	12	Confirmação de ORF

326	c184, c198, c200, c202	fbaB	Fructose-bisphosphate aldolase class 1	Metabolismo de glicose	4.1.2.13	17, 18, 21, 28	11	7	Confirmação de ORF
327	c33	mtlK	Mannitol 2-dehydrogenase	Metabolismo de glicose	1.1.1.67	21	6	-	Confirmação de ORF
328		mtnB	Methylthioribulose-1-phosphate dehydratase	Metabolismo de glicose	4.2.1.109	3	4	-	Confirmação de ORF
329	c287	mtnD	putative 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	Metabolismo de glicose	-	-	3	6	Confirmação de ORF
330	c239	otsB	putative trehalose-phosphatase	Metabolismo de glicose	3.1.3.12	25	4	3	Confirmação de ORF
331		pckA	Phosphoenolpyruvate carboxykinase	Metabolismo de glicose	4.1.1.49	19, 20, 28	2	-	Confirmação de ORF
332		treY	Maltooligosyl trehalose synthase	Metabolismo de glicose	5.4.99.15	-	6	-	Confirmação de ORF
333		atpA	ATP synthase subunit alpha	Metabolismo de pentoses-fostato	3.6.3.14	-	19	3	Confirmação do N-terminal
334	c330	atpC	ATP synthase epsilon chain	Metabolismo de pentoses-fostato	3.6.3.14	-	-	6	Confirmação de ORF
335	c34, c36	atpD	ATP synthase subunit beta	Metabolismo de pentoses-fostato	3.6.3.14	-	50	9	Confirmação do N-terminal
336		atpH	ATP synthase delta chain	Metabolismo de pentoses-fostato	3.6.3.14	-	5	-	Confirmação de ORF
337	c164r, c177, c182	gnd	6-phosphogluconate dehydrogenase	Metabolismo de pentoses-fostato	1.1.1.43	6, 18	45	4	Confirmação do N-terminal
338	c165, c248	pgl	6-phosphogluconolactonase	Metabolismo de pentoses-fostato	3.1.1.31	18	-	8	Confirmação de ORF
339	c241	rpiA	Ribose-5-phosphate isomerase A	Metabolismo de pentoses-fostato	5.3.1.6	18, 28	4	2	Confirmação de ORF
340		tal	Transaldolase	Metabolismo de pentoses-fostato	2.2.1.2	18	33	-	Confirmação do C-terminal

341	c16	tktA	Transketolase	Metabolismo de pentoses-fostato	2.2.1.1	18, 28	34	6	Confirmação de ORF
342	c45	zwf	Glucose-6-phosphate 1-dehydrogenase	Metabolismo de pentoses-fostato	1.1.1.49	6, 18	18	8	Confirmação de ORF
343	c45	zwf	Glucose-6-phosphate 1-dehydrogenase	Metabolismo de pentoses-fostato	1.1.1.49	6, 18	7	-	Confirmação de ORF
344		3_0101	putative D-2-hydroxyacid dehydrogenase	Metabolismo de pequenas moléculas	1.1.1.272	-	38	-	Confirmação de ORF
345		3_0156	putative quinone oxidoreductase	Metabolismo de pequenas moléculas	1.12.5.1	-	2	-	Confirmação de ORF
346		5_0362	putative hydrolase	Metabolismo de pequenas moléculas	-	-	1	-	Confirmação de ORF
347		6_0139	Ribitol 2-dehydrogenase	Metabolismo de pequenas moléculas	1.1.1.56	-	23	-	Confirmação do C-terminal
348		6_0538	Ferredoxin--NADP reductase	Metabolismo de pequenas moléculas	1.18.1.2	-	1	-	Confirmação de ORF
349		6_0612	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase	Metabolismo de pequenas moléculas	4.2.1.59	-	2	-	Confirmação de ORF
350	c163r	6_0811	putative peptidyl-prolyl cis-trans isomerase D	Metabolismo de pequenas moléculas	5.2.1.8	-	11	-	Confirmação de ORF
351		idnD	putative L-idonate 5-dehydrogenase	Metabolismo de pequenas moléculas	1.1.1.264	-	2	-	Confirmação de ORF
352		2_0125	putative short-chain dehydrogenase	Metabolismo intermediário central	-	-	2	-	Confirmação de ORF
353	c63	3_0305	putative glutamate carboxypeptidase protein	Metabolismo intermediário central	-	-	-	20	Confirmação de ORF
354		3_0486	putative Deoxycytidine triphosphate deaminase	Metabolismo intermediário central	3.5.4.13	38	2	-	Confirmação de ORF
355		3_0762	putative phosphoketolase	Metabolismo intermediário central	4.1.2.9	18, 28, 29	24	-	Confirmação de ORF
356		4_00004	putative glycosyl transferase	Metabolismo intermediário central	-	-	2	-	Confirmação de ORF
357	c232	4_00278	putative hydroxyacylglutathione	Metabolismo intermediário central	3.1.2.6	19	1	7	Confirmação de ORF

		hydrolase							
358		4_00334 putative GTP-binding protein	Metabolismo intermediário central	-	-	1	-		Confirmação de ORF
359		5_0003 putative oxidoreductase	Metabolismo intermediário central	-	-	7	-		Confirmação de ORF
360		5_0037 Nitrilase/cyanide hydratase	Metabolismo intermediário central	4.2.1.66	-	1	-		Confirmação de ORF
361	c166	5_0278 dehydrogenase (zinc-binding alcohol dehydrogenase)	Metabolismo intermediário central	-	-	1	-		Confirmação de ORF
362	c154	5_0311 putative thioredoxin protein	Metabolismo intermediário central	-	38	1	8		Confirmação de ORF
363		5_0315 Zinc-type alcohol dehydrogenase-like protein	Metabolismo intermediário central	-	-	1	7		Confirmação de ORF
364	c190	5_0483 Phosphatase	Metabolismo intermediário central	-	-	-	8		Confirmação de ORF
365	c196	5_0484 Phosphatase	Metabolismo intermediário central	-	-	2	7		Confirmação de ORF
366		5_0592 putative isochorismatase family protein	Metabolismo intermediário central	3.3.2.1	-	1	-		Confirmação de ORF
367	c166r	5_0757 probable oxidoreductase (aldo/keto reductase) protein	Metabolismo intermediário central	-	-	-	8		Confirmação de ORF
368	c150	5_0858 putative aminotransferase	Metabolismo intermediário central	-	-	-	9		Confirmação de ORF
369	c150	5_0858 putative aminotransferase	Metabolismo intermediário central	-	-	-	8		Confirmação de ORF
370	c150	5_0858 putative aminotransferase	Metabolismo intermediário central	-	-	-	8		Confirmação de ORF
371	c100	5_1016 Inositol-3-phosphate synthase	Metabolismo intermediário central	5.5.1.4	36	2	13		Confirmação do N-terminal
372		5_1026 putative glycosyl transferase	Metabolismo intermediário central	-	-	-	9		Confirmação de ORF
373		5_1026 putative glycosyl transferase	Metabolismo intermediário central	-	-	-	9		Confirmação de ORF
374		5_1107 Gluconate 5-dehydrogenase	Metabolismo intermediário central	1.1.1.69	-	10	-		Confirmação de ORF

375		5_1179	putative oxidoreductase	Metabolismo intermediário central	-	-	11	-	Confirmação de ORF
376	c118	5_1475	putative GTP-dependent nucleic acid-binding protein engD	Metabolismo intermediário central	-	-	2	17	Confirmação de N e C-terminal
377	c243	5_1490	putative beta-lactamase-like	Metabolismo intermediário central	3.5.2.6	40	3	2	Confirmação de ORF
378		5_1578	Putative NAD(P)H nitroreductase	Metabolismo intermediário central	-	-	1	-	Confirmação de ORF
379		5_1650	Putative aminotransferase	Metabolismo intermediário central	-	-	2	-	Confirmação de ORF
380	c240	6_0131	putative peroxiredoxin	Metabolismo intermediário central	-	-	-	4	Confirmação de ORF
381		6_0156	Glucose 1-dehydrogenase	Metabolismo intermediário central	1.1.1.47	18	1	-	Confirmação de ORF
382	c258	6_0204	5,10-methylenetetrahydrofolate reductase	Metabolismo intermediário central	1.5.1.20	29, 32	-	10	Confirmação de ORF
383		6_0253	putative nucleoside diphosphate hydrolase	Metabolismo intermediário central	3.6.1.6	37, 38	1	-	Confirmação de ORF
384		6_0385	Aminopeptidase	Metabolismo intermediário central	3.4.11.-	-	-	9	Confirmação de ORF
385		6_0395	Glutathione S-transferase	Metabolismo intermediário central	2.5.1.18	6	3	-	Confirmação de ORF
386	c145	6_0470	putative glycosyl transferase, group 1	Metabolismo intermediário central	-	-	1	8	Confirmação de ORF
387		6_0478	Short-chain dehydrogenase	Metabolismo intermediário central	-	-	3	-	Confirmação de ORF
388		6_0803	Dihydrolipoamid acetyltransferase component of pyruvate dehydrogenase complex	Metabolismo intermediário central	R02569	17	5	-	Confirmação de ORF
389	c176, c247	6_0804	Pyruvate dehydrogenase E1 component subunit beta	Metabolismo intermediário central	1.2.4.1	5, 7, 17, 19,	8	5	Confirmação de ORF

390	6_0805	Pyruvate dehydrogenase E1 component subunit alpha	Metabolismo intermediário central	1.2.4.1	22 5, 7, 17, 19, 22	1	-	Confirmação de ORF
391	6_0901	Acetyl-CoA hydrolase	Metabolismo intermediário central	3.1.2.1	19	5	-	Confirmação de ORF
392	c162	accD putative Acetyl-coenzyme A carboxylase carboxyl transferase	Metabolismo intermediário central	-	-	2	-	Confirmação de ORF
393	c104	ackA putative acetate kinase	Metabolismo intermediário central	2.7.2.1	19	13	4	Confirmação de ORF
394	c15	acnA putative Aconitate hydratase	Metabolismo intermediário central	4.2.1.3	20, 27, 31	29	-	Confirmação de ORF
395	c269	acpD Putative acyl carrier protein phosphodiesterase	Metabolismo intermediário central	3.1.4.14	-	1	5	Confirmação do N-terminal
396	c140	argD Acetylornithine aminotransferase	Metabolismo intermediário central	2.6.1.11	4, 5, 11, 13, 14	8	1	Confirmação do N-terminal
397	c82, c83	aroA 3-phosphoshikimate 1-carboxyvinyltransferase	Metabolismo intermediário central	2.5.1.19	4	3	12	Confirmação de ORF
398	cpo	alpha/beta hydrolase, chloride peroxidase	Metabolismo intermediário central	1.11.1.10	-	1	-	Confirmação de ORF
399	c249	dapA Dihydrodipicolinate synthase	Metabolismo intermediário central	-	-	-	10	Confirmação de ORF
400	c197	exoN UTP--glucose-1-phosphate uridylyltransferase	Metabolismo intermediário central	-	-	1	12	Confirmação de ORF
401	c187	fabD Malonyl CoA-acyl carrier protein transacylase	Metabolismo intermediário central	-	-	-	10	Confirmação de ORF
402	c85	gor Glutathione reductase	Metabolismo intermediário central	-	-	-	14	Confirmação de ORF
403	c171	luxA putative Alkanal monooxygenase	Metabolismo intermediário central	1.14.14.3	-	-	12	Confirmação de N e C-terminal

404	c258	metF	5,10-methylenetetrahydrofolate reductase	Metabolismo intermediário central	-	-	-	10	Confirmação de ORF
405	c92	metK	putative S-adenosylmethionine synthetase	Metabolismo intermediário central	2.5.1.6	3, 11	3	12	Confirmação de ORF
406		mod1	NADP-dependent malic enzyme	Metabolismo intermediário central	1.1.1.82	19, 28	3	-	Confirmação de ORF
407		mraW	S-adenosyl-methyltransferase mraW	Metabolismo intermediário central	-	-	1	-	Confirmação de ORF
408		murA	putative UDP-N-acetylglucosamine 1-carboxyvinyltransferase	Metabolismo intermediário central	2.5.1.7	-	1	-	Confirmação de ORF
409	c58	opgG	Glucans biosynthesis protein G	Metabolismo intermediário central	-	-	-	11	Confirmação de ORF
410	c239	otsB	putative trehalose-phosphatase	Metabolismo intermediário central	-	-	-	8	Confirmação de ORF
411	c282	ppa	Inorganic pyrophosphatase	Metabolismo intermediário central	-	-	3	4	Confirmação de ORF
412		purF	putative Amidophosphoribosyltransferase precursor	Metabolismo intermediário central	2.4.2.14	1, 37	1	-	Confirmação de ORF
413	c52	sucB	2-oxoglutarate dehydrogenase E2 component	Metabolismo intermediário central	1.2.1.52	-	15	3	Confirmação de ORF
414		sucC	2-oxoglutarate dehydrogenase E3 component	Metabolismo intermediário central	1.2.4.2	15, 20	4	-	Confirmação do N-terminal
415	c79	thrC	Threonine synthase	Metabolismo intermediário central	-	-	-	7	Confirmação de ORF
416		ydfG	NADP-dependent L-serine/L-allo-threonine dehydrogenase	Metabolismo intermediário central	-	-	2	-	Confirmação de ORF
417	c260	4_00338	Putative carboxymethylenebutenolidase	Outra	3.1.1.45	-	3	-	Confirmação de ORF
418	c227	5_0950	putative enolase-phosphatase	Outra	-	-	-	7	Confirmação de ORF
419		6_01024	Type II/IV secretion system protein	Outra	-	-	3	-	Confirmação de ORF
420		6_0492	Nitrogen assimilation	Outra	-	-	1	-	Confirmação de ORF

		regulatory protein ntrX							
421		infC Translation initiation factor IF-3	Outra	-	-	7	-		Confirmação de ORF
422		nifA Nif-specific regulatory protein	Outra	-	-	2	-		Confirmação de ORF
423		nifD Nitrogenase molybdenum-iron protein alpha chain	Outra	-	-	3	-		Confirmação de ORF
424		nifH Nitrogenase iron protein	Outra	-	-	-	-		Confirmação de ORF
425		nifK Nitrogenase molybdenum-iron protein beta chain	Outra	-	-	1	-		Confirmação de ORF
426		suhB Inositol-1-monophosphatase	Outra	3.1.3.25	36	1	-		Confirmação de ORF
427		exsB putative succinoglycan	Polissacarídeos	-	-	1	-		Confirmação de ORF
428	c58	opgG Glucans biosynthesis protein G	Polissacarídeos	-	-	-	9		Confirmação de ORF
429		6_01050 Chemoreceptor mcpA (Methyl-accepting chemotaxis protein)	Processos celulares	-	-	1	-		Confirmação de ORF
430		6_0522 Chaperone protein htpG	Processos celulares	-	-	15	-		Confirmação de ORF
431		cheW putative chemotaxis protein cheW	Processos celulares	-	-	1	-		Confirmação de ORF
432		dnaJ Chaperone protein dnaJ	Processos celulares	-	-	1	-		Confirmação de ORF
433		dnaK putative Chaperone protein dnaK	Processos celulares	-	-	57	-		Confirmação de ORF
434		grpE putative chaperone binding	Processos celulares	-	-	1	-		Confirmação de ORF
435		mcpA Chemoreceptor mcpA	Processos celulares	-	-	7	-		Confirmação de ORF
436		surA peptidyl-prolyl cis-trans isomerase SurA	Processos celulares	5.2.1.8	-	1	11		Confirmação de ORF
437		2_0239 putative TonB-dependent receptor	Proteínas de transporte e ligação	-	-	8	-		Confirmação de ORF
438		3_0472 putative TonB-dependent receptor	Proteínas de transporte e ligação	-	-	1	-		Confirmação de ORF
439		3_0534 putative TonB-dependent Receptor protein	Proteínas de transporte e ligação	-	-	15	-		Confirmação de ORF
440		4_00009 ABC transporter ATP-binding protein	Proteínas de transporte e ligação	-	-	21	-		Confirmação de ORF
441		4_00291 putative ABC transporter	Proteínas de transporte e ligação	-	-	1	-		Confirmação de ORF

442	4_00542	putative thiamine transport ATP-binding protein	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
443	5_0894	TonB-dependent receptor	Proteínas de transporte e ligação	-	-	4	-	Confirmação de ORF
444	5_1075	TonB-dependent receptor	Proteínas de transporte e ligação	-	-	8	-	Confirmação de ORF
445	6_0119	Fucose operon fucU protein	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
446	c352	6_0124 D-ribose-binding periplasmic protein precursor	Proteínas de transporte e ligação	-	-	3	3	Confirmação de ORF
447	6_0125	L-fucose isomerase	Proteínas de transporte e ligação	5.3.1.25	21	-	3	Confirmação de ORF
448	6_0144	putative sugar uptake ABC transporter ATP-binding protein	Proteínas de transporte e ligação	-	-	2	-	Confirmação de ORF
449	6_0237	D-xylose-binding periplasmic protein precursor	Proteínas de transporte e ligação	-	-	2	-	Confirmação de ORF
450	c183	6_0687 putative periplasmic binding protein	Proteínas de transporte e ligação	-	-	5	-	Confirmação de ORF
451	6_083	TonB-dependent receptor	Proteínas de transporte e ligação	-	-	2	-	Confirmação de ORF
452	6_0995	ABC transporter ATP-binding protein uup	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
453	exbB	putative Biopolymer transport exbB protein	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
454	c193	murB UDP-N-acetylenolpyruvoylglucosamin e reductase	Proteínas de transporte e ligação	-	-	-	5	Confirmação de ORF
455	c235, c236	PBP putative phosphate-binding periplasmic protein precursor	Proteínas de transporte e ligação	-	-	1	9	Confirmação de ORF
456	phoU	Phosphate transport system protein	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
457	phoU	putative phosphate transport system protein phoU	Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
458	sufC	FeS assembly protein SufC	Proteínas de transporte	-	-	1	-	Confirmação de ORF

459		tolB	Protein tolB precursor	e ligação Proteínas de transporte e ligação	-	-	1	-	Confirmação de ORF
460	c344	3_0419	putative ROS/MUCR transcriptional regulator protein	Regulação	-	-	-	4	Confirmação de ORF
461		6_0847	BolA-like protein	Regulação	-	-	1	-	Confirmação de ORF
462	c291	6_0979	N utilization substance protein B homolog	Regulação	-	-	1	6	Confirmação do N-terminal
463	c304	6_0437	Single-stranded DNA-binding protein	Replicação, reparo e modificação de DNA	-	-	10	4	Confirmação de ORF
464		6_0463	Protein recA (Recombinase A)	Replicação, reparo e modificação de DNA	-	-	8	-	Confirmação de ORF
465		6_0963	DNA polymerase III subunit beta	Replicação, reparo e modificação de DNA	2.7.7.7	37, 38	6	-	Confirmação de ORF
466		hup	DNA-binding protein HU	Replicação, reparo e modificação de DNA	-	-	1	-	Confirmação de ORF
467		recN	DNA repair protein recN	Replicação, reparo e modificação de DNA	-	-	1	-	Confirmação de ORF
468	c123	rpoA	DNA-directed RNA polymerase alpha chain	Replicação, reparo e modificação de DNA	2.7.7.6	40, 37, 38	11	10	Confirmação de ORF
469		rpoB	DNA-directed RNA polymerase beta chain	Replicação, reparo e modificação de DNA	2.7.7.6	40, 37, 38	13	-	Confirmação de ORF
470		rpoC	DNA-directed RNA polymerase beta' chain	Replicação, reparo e modificação de DNA	2.7.7.6	40, 37, 38	4	-	Confirmação de ORF
471	c140	ruvB	Holliday junction ATP-dependent DNA helicase ruvB	Replicação, reparo e modificação de DNA	-	-	-	8	Confirmação de ORF
472		4_00369	putative universal stress protein	Respostas de proteção	-	-	5	-	Confirmação do C-terminal
473	c185	4_00623	putative nucleoside diphosphate	Respostas de proteção	-	-	-	9	Confirmação de ORF
474	c208	5_0691	putative bacteriocin protein	Respostas de proteção	-	-	7	2	Confirmação do N-

475	c283, c285	5_0935	Peroxiredoxin	Respostas de proteção	1.11.1.15	-	12	8	terminal Confirmação de N e C-terminal
476	c299	6_0587	Superoxide dismutase [Fe]	Respostas de proteção	1.15.1.1	-	17	2	Confirmação de ORF
477		6_0635	putative organic solvent tolerance protein	Respostas de proteção	-	-	1	-	Confirmação de ORF
478		emrA	Multidrug resistance protein A	Respostas de proteção	-	-	1	-	Confirmação de ORF
479		pqiB	Paraquat-inducible protein B	Respostas de proteção	-	-	1	-	Confirmação de ORF
480		gatA	Glutamyl-tRNA(Gln) amidotransferase subunit A	Síntese de aminoacil tRNA, modificação de tRNA	6.3.5.7	1	3	-	Confirmação de ORF
481		gatB	Aspartyl/glutamyl- tRNA(Asn/Gln) amidotransferase subunit	Síntese de aminoacil tRNA, modificação de tRNA	-	-	4	-	Confirmação de ORF
482	c80	gltX	Glutamyl-tRNA synthetase I	Síntese de aminoacil tRNA, modificação de tRNA	6.1.1.17	1, 33, 39	6	8	Confirmação de ORF
483		prfB	putative peptide chain release factor 2	Síntese de aminoacil tRNA, modificação de tRNA	-	-	1	-	Confirmação de ORF
484		proS	Prolyl-tRNA synthetase	Síntese de aminoacil tRNA, modificação de tRNA	6.1.1.15	9, 39	5	-	Confirmação de ORF
485	c3, c4	fusA	Elongation factor G	Síntese e modificação de proteínas	-	-	19	19	Confirmação de ORF
486	c158, c161	tst	Elongation factor Ts (EF-Ts)	Síntese e modificação de proteínas	-	-	15	10	Confirmação de ORF
487	c96	tufB	Elongation factor Tu	Síntese e modificação de proteínas	-	-	-	33	Confirmação de ORF
488	c5	typA	GTP-binding protein TypA/BipA	Síntese e modificação de proteínas	-	-	18	11	Confirmação de ORF
489		6_0559	30S ribosomal protein S4	Síntese e modificação de proteínas ribossomais	-	-	6	-	Confirmação do C- terminal
490	c287,	6_0597	DNA-directed RNA	Síntese e modificação	-	-	4	1	Confirmação de ORF

	c289		polymerase omega chain	de proteínas ribossomais					
491	c114	6_0604	Ribosome recycling factor	Síntese e modificação de proteínas ribossomais	-	-	5	2	Confirmação de ORF
492		rplA	50S ribosomal protein L1	Síntese e modificação de proteínas ribossomais	-	-	2	-	Confirmação de ORF
493		rplC	50S ribosomal protein L3	Síntese e modificação de proteínas ribossomais	-	-	9	-	Confirmação de ORF
494		rplC	50S ribosomal protein L3	Síntese e modificação de proteínas ribossomais	-	-	6	-	Confirmação de ORF
495		rplD	50S ribosomal protein L4	Síntese e modificação de proteínas ribossomais	-	-	3	-	Confirmação do N-terminal
496		rplE	50S ribosomal protein L5	Síntese e modificação de proteínas ribossomais	-	-	7	-	Confirmação de ORF
497		rplF	50S ribosomal protein L6	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF
498	c268, c285	rplI	50S ribosomal protein L9	Síntese e modificação de proteínas ribossomais	-	-	25	4	Confirmação de ORF
499		rplJ	50S ribosomal protein L10	Síntese e modificação de proteínas ribossomais	-	-	6	-	Confirmação de ORF
500		rplO	50S ribosomal protein L15	Síntese e modificação de proteínas ribossomais	-	-	4	-	Confirmação de ORF
501		rplV	/systematic_id=rpsC	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF

502		rplX	50S ribosomal protein L24	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF
503	c113	rplY	50S ribosomal protein L25	Síntese e modificação de proteínas ribossomais	-	-	12	3	Confirmação de ORF
504	c30	rpsA	30S ribosomal protein S1	Síntese e modificação de proteínas ribossomais	-	-	87	3	Confirmação de ORF
505		rpsB	30S ribosomal protein S2	Síntese e modificação de proteínas ribossomais	-	-	10	-	Confirmação de ORF
506		rpsE	30S ribosomal protein S5	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF
507	c319	rpsF	30S ribosomal protein S6	Síntese e modificação de proteínas ribossomais	-	-	1	8	Confirmação do N-terminal
508		rpsI	30S ribosomal protein S9	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF
509		rpsS	30S ribosomal protein S19	Síntese e modificação de proteínas ribossomais	-	-	1	-	Confirmação de ORF
510		3_0366	putative translation initiation factor IF-2	Síntese e modificação de RNA, transcrição do DNA	-	-	2	-	Confirmação de ORF
511	c7	4_00690	putative RNA-metabolising metallo-beta-lactamase protein	Síntese e modificação de RNA, transcrição do DNA	-	-	10	11	Confirmação de ORF
512		6_0632	putative sigma(54) modulation protein	Síntese e modificação de RNA, transcrição do DNA	-	-	5	-	Confirmação do C-terminal
513		greA	Transcription elongation factor greA	Síntese e modificação de RNA, transcrição	-	-	1	-	Confirmação de ORF

514		nusA	putative transcription elongation protein nusA	do DNA Síntese e modificação de RNA, transcrição do DNA	-	-	15	-	Confirmação de ORF
515	c9	pnp	putative polyribonucleotide nucleotidyltransferase	Síntese e modificação de RNA, transcrição do DNA	2.7.7.8	37, 38	27	13	Confirmação de N e C-terminal
516		rho	putative Transcription termination factor rho	Síntese e modificação de RNA, transcrição do DNA	-	-	12	-	Confirmação de ORF
517		3_0011	putative guanylate kinase	Tradução e modificação de proteínas	2.7.4.8	37	2	-	Confirmação de ORF
518		5_1266	Peptidylprolyl isomerase	Tradução e modificação de proteínas	-	-	1	-	Confirmação de ORF
519	c37	6_0134	60 kDa chaperonin	Tradução e modificação de proteínas	-	-	6	7	Confirmação do C-terminal
520		6_0523	Putative translation initiation factor	Tradução e modificação de proteínas	-	-	3	-	Confirmação de ORF
521		6_0568	putative Thiaminase-2	Tradução e modificação de proteínas	3.5.99.2	-	1	-	Confirmação de ORF
522		6_0697	10 kDa chaperonin	Tradução e modificação de proteínas	-	-	1	-	Confirmação de ORF
523	c41, c42, c43	6_0698	60 kDa chaperonin	Tradução e modificação de proteínas	-	-	75	1	Confirmação de ORF
524		6_0974	Protein-L-isoaspartate O-methyltransferase	Tradução e modificação de proteínas	2.1.1.77	-	2	-	Confirmação de ORF
525		efp	Elongation factor P	Tradução e	-	-	2	-	Confirmação de ORF

526		hslO	putative chaperone protein	modificação de proteínas Tradução e modificação de proteínas	-	-	1	-	Confirmação de ORF
527	c32	tig	Trigger factor (TF)	Tradução e modificação de proteínas	-	-	16	5	Confirmação de ORF
528		5_1103	transcription antitermination protein NusG	Tradução e modificação de proteínas	-	-	4	-	Confirmação de ORF
529		6_0575	putative transcriptional regulator, IclR family/regucalcin	Trandução de sinais	-	-	4	-	Confirmação de ORF
530		6_0887	transcriptional regulator, CarD family	Trandução de sinais	-	-	6	-	Confirmação de ORF
531	c275	3_0506	putative two-component response regulator	Transdução de sinais	-	-	11	3	Confirmação do N-terminal
532		4_00103	putative two component Transcriptional regulator	Transdução de sinais	-	-	4	-	Confirmação de ORF
533		chvI	putative Transcriptional regulatory protein	Transdução de sinais	-	-	1	-	Confirmação de ORF
534		LysR	putative transcriptional regulator LysR	Transdução de sinais	-	-	6	-	Confirmação de ORF
535		LysR	putative transcriptional regulator protein	Transdução de sinais	-	-	3	-	Confirmação do C-terminal
536	c131	pheS	Phenylalanyl-tRNA synthetase alpha chain	Transferência de informação	-	-	-	15	Confirmação de ORF
537		ruvB	Holliday junction ATP-dependent DNA helicase ruvB	Transferência de informação	-	-	-	7	Confirmação de ORF
538		2_0501	putative electron transport protein SCO1/SenC	Transporte de elétrons	-	-	2	-	Confirmação de ORF
539		3_0105	putative oxidoreductase protein	Transporte de elétrons	-	-	29	-	Confirmação de frameshift
540		3_0374	putative 2-nitropropane dioxygenase	Transporte de elétrons	1.13.11.3 2	30	2	-	Confirmação de ORF

541	4_00308	putative aldehyde dehydrogenase protein	Transporte de elétrons	-	-	1	-	Confirmação de ORF
542	4_00332	putative pyrrolo-quinoline quinone	Transporte de elétrons	C00113	10	2	-	Confirmação de ORF
543	5_0390	flavin oxidoreductase	Transporte de elétrons	-	-	1	-	Confirmação de ORF
544	5_0855	putative quinone oxidoreductase	Transporte de elétrons	-	-	2	-	Confirmação de ORF
545	5_0964	Putative NAD(P)H nitroreductase	Transporte de elétrons	-	-	2	-	Confirmação de ORF
546	6_0707	Alcohol dehydrogenase [acceptor] precursor	Transporte de elétrons	1.1.99.8	17, 22, 29	14	14	Confirmação de ORF
547	6_0711	Ubiquinol oxidase subunit 2 precursor	Transporte de elétrons	-	-	1	-	Confirmação de ORF
548	c205	etfA putative Electron transfer flavoprotein subunit alpha	Transporte de elétrons	-	-	5	9	Confirmação de ORF
549		etfB putative electron transfer flavoprotein subunit beta	Transporte de elétrons	-	-	10	-	Confirmação de ORF
550		etfD putative electron transfer flavoprotein-ubiquinone oxidoreductase	Transporte de elétrons	-	-	4	-	Confirmação de ORF
551		fdhH putative formate dehydrogenase alpha subunit	Transporte de elétrons	1.2.1.2	27, 29, 40	5	-	Confirmação de ORF
552		fixA Protein fixA	Transporte de elétrons	-	-	2	-	Confirmação de ORF
553		namA NADPH dehydrogenase	Transporte de elétrons	1.6.99.1	-	1	-	Confirmação de ORF
554		nouC NADH-quinone oxidoreductase chain C	Transporte de elétrons	1.6.5.2	35	5	-	Confirmação de ORF
555		nouE NADH-quinone oxidoreductase chain E	Transporte de elétrons	1.6.5.2	35	1	-	Confirmação de ORF
556		nouF NADH-quinone oxidoreductase chain F	Transporte de elétrons	1.6.5.2	35	2	-	Confirmação de ORF
557		nouG NADH-quinone oxidoreductase chain G	Transporte de elétrons	1.6.5.2	35	10	-	Confirmação de ORF
558		nouI NADH-quinone	Transporte de elétrons	1.6.5.2	35	1	-	Confirmação de ORF

559		nuoB	oxidoreductase chain I NADH-quinone oxidoreductase chain B	Transporte de elétrons	1.6.5.2	35	2	-	Confirmação de ORF
560	c284	rutE	Putative NADH dehydrogenase/NAD(P)H nitroreductase rutE	Transporte de elétrons	-	-	-	4	Confirmação de ORF
561		sdhA	putative succinate dehydrogenase flavoprotein subunit	Transporte de elétrons	-	-	21	-	Confirmação de ORF
562		sdhB	putative succinate dehydrogenase iron-sulfur protein	Transporte de elétrons	-	-	3	-	Confirmação de ORF
563		ubiE	Ubiquinone/menaquinone biosynthesis methyltransferase ubi	Transporte de elétrons	-	-	-	8	Confirmação de ORF
564		6_0524	Aspartyl-tRNA synthetase	tRNA	6.1.1.12	5, 39	17	-	Confirmação de ORF
565		6_0530	Arginyl-tRNA synthetase	tRNA	6.1.1.19	9, 39	3	-	Confirmação de ORF
566		6_0536	Seryl-tRNA synthetase	tRNA	6.1.1.11	2, 39	5	-	Confirmação de ORF
567		6_0567	Alanyl-tRNA synthetase (Alanine--tRNA ligase) (AlaRS)	tRNA	6.1.1.7	5, 39	7	-	Confirmação de ORF
568		6_0624	Methionyl-tRNA formyltransferase	tRNA	2.1.2.9	3, 32, 39	1	-	Confirmação de ORF
569		aatA	Aspartate aminotransferase	tRNA	2.6.1.1	1, 9, 10, 28	4	-	Confirmação de ORF
570	c139	pheS	Phenylalanyl-tRNA synthetase alpha chain	tRNA	6.1.1.20	4, 39	-	16	Confirmação de ORF
571		pheT	Phenylalanyl-tRNA synthetase beta chain	tRNA	6.1.1.20	4, 39	1	-	Confirmação de ORF
572	c72	serS	Seryl-tRNA synthetase	tRNA	6.1.1.11	-	-	5	Confirmação de ORF
573		thrS	Threonyl-tRNA synthetase	tRNA	6.1.1.3	2, 39	8	-	Confirmação do C-terminal

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](#)