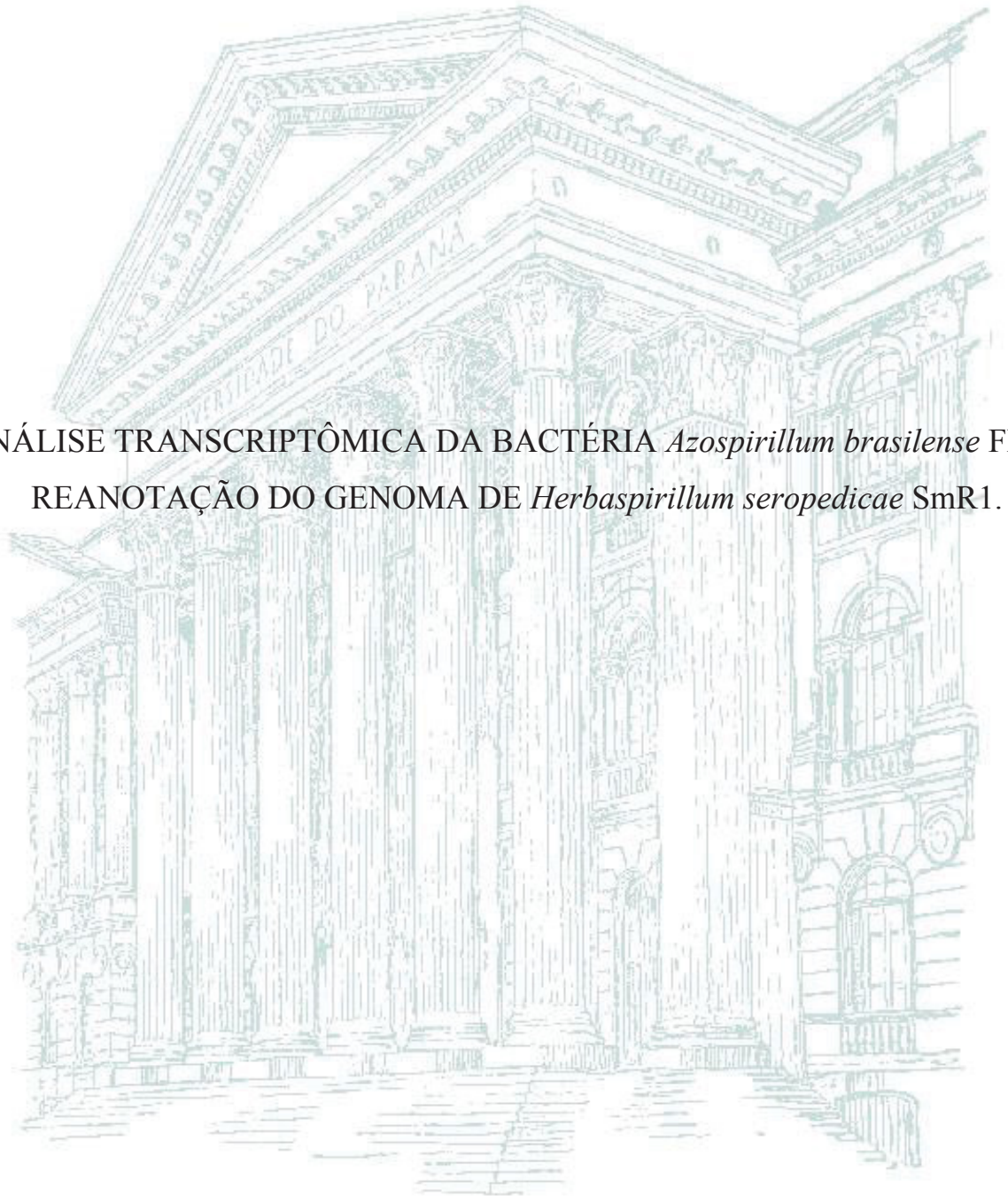
UNIVERSIDADE FEDERAL DO PARANÁ

HELBA CIRINO DE SOUZA BARBOSA



ANÁLISE TRANSCRIPTÔMICA DA BACTÉRIA *Azospirillum brasilense* FP2 E REANOTAÇÃO DO GENOMA DE *Herbaspirillum seropedicae* SmR1.

CURITIBA-PR

2018

HELBA CIRINO DE SOUZA BARBOSA

ANÁLISE TRANSCRIPTÔMICA DA BACTÉRIA *Azospirillum brasilense* FP2 E REANOTAÇÃO DO GENOMA DE *Herbaspirillum seropedicae* SmR1.

Tese de doutorado apresentado ao Programa de Pós-Graduação em Ciências-Bioquímica, Departamento de Bioquímica e Biologia Molecular, Setor de Ciências Biológicas, Universidade Federal do Paraná.

Orientadora: Prof ª. Rose Adele Monteiro.
Co-orientador: Prof. Emanuel Maltempi de Souza

CURITIBA-PR
2018

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Telma Terezinha Stresser de Assis –CRB/9-944)

Barbosa, Helba Cirino de Souza
Análise transcriptomica da bactéria *Azospirillum brasiliense* FP2 e reanotação do genoma de *Herbaspirillum seropedicae* SmR1. / Helba Cirino de Souza Barbosa. – Curitiba, 2018.
177 p.: il. ; 30cm.

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

1. Nitrogênio - Fixação. 2. Bactérias. 3. Azospirillum. 4. Herbaspirillum. 5. Transcriptoma. I. Título. II. Monteiro, Rose Adele. III. Souza, Emanuel Maltempi de. IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Ciências - Bioquímica.

CDD (20. ed.) 574.873283

MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS
(BIOQUÍMICA)

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS (BIOQUÍMICA) da Universidade Federal do Paraná foram convocados para realizar a arguição da tese de Doutorado de **HELBA CIRINO DE SOUZA BARBOSA** intitulada: **Análise transcriptômica da bactéria *Azospirillum brasilense* FP2 e Reanotação do genoma de *Herbaspirillum seropedicae* SmR1** , após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação no rito de defesa.

A outorga do título de doutor está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 25 de Junho de 2018.

ROSE ADELE MONTEIRO
Presidente da Banca Examinadora

FABIO PASSETTI
Avaliador Externo

JERONIZA NUNES MARCHAUKOSKI
Avaliador Externo

MAURO ANTONIO ALVES CASTRO
Avaliador Externo

LEONARDO MAGALHÃES CRUZ
Avaliador Interno

AGRADECIMENTOS

-Ao Omar Cirino de Souza (papai), e à Edna Sousa Barbosa (mamãe) por sempre terem me incentivado nos estudos, dando todo o apoio necessário;

-Aos professores Dra. Rose Adele Monteiro, Dr. Emanuel Maltempi de Souza pela oportunidade e orientações;

- Aos professores Dr. Leonardo Magalhães Cruz e Marcelo Müller dos Santos por serem a banca interna oferecendo correções e orientações;

- À professora Dra. Maria Berenice Reynaud Steffens;

- À Dra. Michele Tadra, ao Dr. Rodrigo Luiz Cardoso;

- Aos colegas do laboratório de Bioinformática.

- À amiga Susana Gobetti; e à amiga Mariana Kirsten, por se fazer presente em minha vida, mesmo distante;

-Aos amigos-irmãos CEBUDV;

-À Capes e CNPq.

RESUMO

Herbaspirillum seropedicae e *Azospirillum brasilense* são bactérias fixadoras de Nitrogênio, encontradas associadas à plantas de interesse agrícola e econômico para o país, como milho, arroz e trigo. São bactérias promotoras do crescimento vegetal, sendo utilizadas como biofertilizantes. Na primeira parte deste trabalho *A. brasilense* FP2 teve seu genoma montado novamente, reanotado e utilizado para o mapeamento dos resultados de RNA-Seq de *Azospirillum brasilense* cultivado em condições de fixação de nitrogênio. Os genes expressos diferencialmente foram identificados. Com esses dados obteve-se o panorama metabólico de FP2 crescida em condições de fixação de Nitrogênio. Foi possível comparar os resultados de *A. brasilense* com outras espécies fixadoras de Nitrogênio, em que se destaca a repressão da Montagem Flagelar em *A. brasilense.* Na segunda parte deste trabalho, o genoma de *H. seropedicae* SmR1 foi reanotado, utilizando dados de RNA-Seq. Identificamos 309 genes novos e atualizamos as classes funcionais dos genes. Também foi possível identificar 901 operons no genoma.

Palavras-Chave: Fixação biológica de nitrogênio, bactérias diazotróficas, *H. seropedicae*, *A. brasilense*, RNA-seq.

ABSTRACT

Azospirillum brasilense and Herbaspirillum seropedicae are diazotrophic bacteria, living associated with plants of agricultural and economic interest to the country, such as corn, rice and wheat. They are bacteria that promote plant growth and are used as biofertilizers. In the first part of this work *A. brasilense* FP2 had its genome re-assembled, reanoted and used for the mapping of the RNA-Seq results of *Azospirillum brasilense* cultivated under nitrogen fixation conditions. Differentially expressed genes were identified. With these data we obtained the metabolic panorama of FP2 grown under Nitrogen fixation conditions. It was possible to compare the results of *A. brasilense* with other Nitrogen fixing species, in which the repression of the Flagellar Assembly in *A. brasilense* is highlighted. In the second part of this work, the genome of *H. seropedicae* SmR1 was reanoted, using RNA-Seq data. We identified 309 new genes and upgraded the functional classes of the genes. It was also possible to identify 901 operons in the genome.

Key words: diazotrophic bacteria, *H. seropedicae*, *A. brasilense*, RNA-seq.

LISTA DE FIGURAS

FIGURA 01: REPRESENTAÇÃO DOS SEIS REPLICONS DE *A. brasilense* FP2....13

FIGURA 02: FLUXOGRAMA DAS ETAPAS DO PROJETO.....21

FIGURA 03: FLUXO DE TRABALHO PARA A MONTAGEM.....34

FIGURA 04: GRÁFICO *DOTPOT* COMPARANDO AS MONTAGENS.....36

FIGURA 05: COMPARAÇÃO DA MONTAGEM *DE NOVO* COM O CROMOSSOMO DO *A. brasilense* SP7.....37

FIGURA 06: GRÁFICOS BOX WHISKER DE ANÁLISE DE QUALIDADE.....41

FIGURA 07: ANÁLISE DE COMPONENTE PRINCIPAL (PCA).....42

FIGURA 08: DIAGRAMA DE VENN COM A QUANTIDADE DE GENES D.E.....44

FIGURA 09: PANORAMA DAS VIAS METABÓLICAS NA BACTÉRIA *A. brasilense* FP2.....45

FIGURA 10: VALORES DE LOG2-FOLD-CHANGE DOS GENES *nif*.....45

FIGURA 11: COMPARAÇÃO DOS VALORES DE LOG2-FOLD-CHANGE DOS GENES *nif*.....46

FIGURA 12: VALORES DE LOG2 FOLD-CHANGE DOS GENES PARA PROTEÍNAS RIBOSSOMAIS.....48

FIGURA 13: COMPARAÇÃO DOS VALORES DE *FOLD-CHANGE* DOS GENES DO METABOLISMO ENERGÉTICO.....50

FIGURA 14: GENES DA VIA DA MONTAGEM FLAGELAR QUE ESTÃO REPRIMIDOS.....51

FIGURA 15: ESTRUTURA DA MONTAGEM FLAGELAR.....52

FIGURA 16: EXEMPLOS DE GENES ADICIONADOS AO GENOMA.....61

FIGURA 17: EXEMPLOS DE ALTERAÇÕES FEITAS NA ANOTAÇÃO DO GENOMA.....62

FIGURA 18: EXEMPLO DE UM GENE REMOVIDO DO GENOMA.....62

FIGURA 19: RECLASSIFICAÇÃO DOS GENES “S - FUNÇÃO DESCONHECIDA”.....66

FIGURA 20: EXEMPLO DE OPERON COMPOSTO PELOS TRÊS GENES.....68

FIGURA 21: EXEMPLO DE AMPLIAÇÃO DE UM OPERON APÓS REANOTAÇÃO.....69

FIGURA 22: OPERON *livKHMGF*.....69

FIGURA 23: MODELO DO OPERON *livKHMGF* EM *Escherichia coli*.....70

LISTA DE TABELAS

TABELA 01: DADOS DE MAPEAMENTO DA MONTAGEM 2.....35

TABELA 02: TESTES DE MAPEAMENTOS REALIZADOS COM O TRANSCRIPTOMA DE *A. brasilense* FP2.....39

TABELA 03: GENES DIFERENCIALMENTE EXPRESSOS DE ACORDO COM TRÊS PROGRAMAS DE ANÁLISE ESTATÍSTICA.....43

TABELA 04: REGULADORES TRANSCRICIONAIS ATIVADOS EM *A. brasilense* FP2 EM CONDIÇÃO DE FIXAÇÃO BIOLÓGICA DE NITROGÊNIO.....54

TABELA 05: GENES ATIVADOS NO TRANSCRIPTOMA DURANTE A FIXAÇÃO BIOLÓGICA DE NITROGÊNIO CUJAS PROTEÍNAS FORAM IDENTIFICADAS NO PROTEOMA.....56

TABELA 06: NOVOS GENES E OPERONS IDENTIFICADOS PARA CINCO ESPÉCIES DE BACTÉRIAS, COM BASE EM DADOS DE RNA-SEQ.....58

TABELA 07: NÚMERO DE NOVOS GENES ANOTADOS PELO NCBI, PELO HGF E PELO GLIMMER.....63

TABELA 08: DADOS DA ANOTAÇÃO ORIGINAL (2011) DO GENOMA DE *H. seropedicae* SMR1 E DA RE-ANOTAÇÃO.....64

TABELA 09: CLASSIFICAÇÃO FUNCIONAL DOS GENES DO GENOMA DA *H. seropedicae* SMR1 DAS TRÊS VERSÕES DE ANOTAÇÃO.....65

TABELA 10: FREQUÊNCIA DA QUANTIDADE DE GENES POR TAMANHO DE OPERON.....71

SUMÁRIO

1. Identificação e caracterização do problema 12
1.1. *Herbaspirillum seropedicae* 12
1.2. *Azospirillum brasilense* 14
1.3. Controle do metabolismo geral de Nitrogênio pelo sistema Ntr 16
1.4. Controle da expressão do complexo enzimático nitrogenase 18
1.5. Controle da atividade do complexo enzimático nitrogenase 18
2. Objetivo Geral 20
2.1. Objetivos Específicos 20
3. Materias e Métodos 21
3.1. Descrição das Condições de cultivo para análise de expressão diferencial em condição limitante de amônio de *Azospirillum brasilense* FP2 24
3.2. *Softwares* 25
3.2.1 DOOR 2.0 (a Database for prOkaryotic OpeRons – Banco de Dados para Operons em Procariotos.) 25
3.2.2 Artemis 16.0 26
3.2.3 CLC DNA Workbench 7.5.5 26
3.2.4 Samtools 0.1.19 26
3.2.5 SHRiMP 2.2.3 (SHort Read Mapping Package) 27
3.2.6 KEGG e KAAS 27
3.2.8 Quality Assessment 0.5 28
3.2.9 FastQC 0.11.5 28
3.2.10 Rockhopper 2.0.3 29
3.2.15. HGF 31
3.2.16. SILA 31
3.2.17. BLAST 32
3.2.18. FGAP 32
3.2.19. G-FINISHER 32
3.2.19. GLIMMER 3.02 33
4. Resultados e Discussões 34
4.1 Transcriptoma do *Azospirillum brasilense* FP2 crescido em condições de fixação de Nitrogênio 34
4.1.1 Remontagem do Genoma da Bactéria *Azospirillum brasilense* FP2. 34
4.1.2. Teste de trimagem e de mapeamento para os transcriptomas de *Azospirillum brasilense* FP2 41

4.1.3. Análise de qualidade das Sequências de RNA 42
4.1.4. Análise de Correlação entre as amostras de RNA-Seq de *Azospirillum brasilense* FP2 43
4.1.5. Identificação dos genes diferencialmente expressos por *A. brasilense* crescido em condições de Fixação de Nitrogênio. 45
4.1.6 Genes envolvidos no Metabolismo de Nitrogênio em *A. brasilense* FP2 47
4.1.7 Genes que codificam para proteínas ribossomais de *A. brasilense* 50
4.1.8 Metabolismo energético de *A. brasilense* crescido em condições de Fixação de Nitrogênio 51
4.1.9 Genes Flagelares de *Azospirillum brasilense* 52
4.1.10 Proteínas Reguladoras de transcrição de *Azospirillum brasilense* 56
4.1.11 Comparação dos dados de transcriptoma e proteoma de *A. brasilense*..... 58
4.2. Re-anotação do Genoma da *Herbaspirillum seropedicae* SmR1 60
4.2.1 Novos genes do Genoma da *Herbaspirillum seropedicae* SmR1 62
4.2.2 Reclassificação funcional do Genoma da *H. seropedicae* SmR1 66
4.2.3 Identificação de operons no Genoma da *H. seropedicae* SmR1 70
Conclusões 74
Referências 75
Anexo 01: Lista de genes reprimidos da *A. brasilense* FP2 em condição de Fixação Biológica de Nitrogênio. 91
Anexo 02: lista dos genes ativados de *A. brasilense* FP2 em condição de Fixação Biológica de Nitrôgenio. 126
Anexo 03: Comparação dos genes diferencialmete expressos da *Azospirillum brasilense* FP2 com os genes diferencialmente expressos de *Azoarcus* sp. BH72, *Azotobacter vinelandii* e *Paenibacillus* sp. WLY78 em condições de Fixação Biológica de Nitrogênio. 136
Anexo 04: Comparação dos dados de RNA-Seq com dados de proteômica da *A. brasilense* FP2. 140
Anexo 05: Tabelas com a relação dos novos genes adicionados ao genoma de *H. seropedicae* SMR1. 146
Anexo 06: Operons preditos pelo *software* Rockhopper ou pelo Banco de Dados DOOR e corrigidos após reanotação 154
Anexo 07: *Script* em Perl para trimar os arquivos de qualidade. 176
Anexo 08: *Script* em Perl para trimar os arquivos csfasta. 177

1. Identificação e caracterização do problema

A interação entre diazotrofos, organismos fixadores de nitrogênio, e a planta resulta muitas vezes em um efeito promotor de crescimento vegetal. Por exemplo, durante a interação entre cana de açúcar e bactérias diazotróficas em associações naturais, a fixação de nitrogênio pode contribuir significativamente para o balanço nitrogenado da planta associada (DOBEREINER, 1992). Este projeto tem como objetivo contribuir para o entendimento das vias metabólicas que são ativadas ou reprimidas quando *A. brasilense* é cultivado em condições de limitação de nitrogênio. O entendimento do metabolismo do nitrogênio permitirá definir o que acontece durante a fixação de nitrogênio e como este estado metabólico pode influenciar na interação da bactéria com a planta. Melhorar o entendimento das vias de sinalização que regulam o metabolismo de nitrogênio da bactéria e da planta e integrar esses fatores essenciais para o desenvolvimento da agricultura de alta produtividade e ambientalmente sustentável.

1.1. *Herbaspirillum seropedicae*

Herbaspirillum seropedicae é uma bactéria vibrióide, gram-negativa, da classe Betaproteobacteria. É um diazotrófico que se associa com gramíneas, como milho, sorgo, cana de açúcar e arroz (BALDANI *et al.*, 1986; OLIVARES *et al*., 1996) e também com plantas tropicais como bananeira e abacaxizeiro (CRUZ *et al*., 2001). É encontrada em raízes, caules e folhas de gramíneas, não sobrevivendo bem em solo natural, principalmente quando estão presentes outros microrganismos, sendo então considerado endofítico obrigatório (BALDANI *et al*., 1997).

Em 2011, *H. seropedicae* SmR1 teve seu genoma completamente sequenciado e anotado pelo programa Genoma do Paraná (GENOPAR) (PEDROSA *et al*., 2011), utilizando sequenciamento automático Sanger e a tecnologia 454 FLX Titanium Roche Pirosequenciamento. Assim foi exposta a maioria dos genes presentes nessa bactéria.

Herbaspirillum seropedicae SmR1 é um mutante espontâneo da estirpe Z78 (BALDANI *et al.*, 1986), resistente a estreptomicina. O genoma consiste em um único cromossomo circular de 5.513.887 pb, com 63,4% de conteúdo GC. Contem 4.735

ORFs, sendo 3.108 genes codificadores de proteínas, com função atribuída e 1.130 com função desconhecida, cobrindo 88,3% do genoma. Sequências codificadoras para 55 tRNA e 3 operons de rRNA (PEDROSA *et al*., 2011).

A anotação do genoma foi feita com uma plataforma de anotação automática com Glimmer 2 e com *softwares* BLAST (ALTSCHUL *et al.*, 1997). Prováveis funções de produto de tradução de ORFs foram inferidas usando o pacote do BLAST (ALTSCHUL *et al.*, 1997), COG (TATUSOV *et al.*, 2003), KEGG (KANEHISA *et al*., 2006) e pFAM (FINN *et al.*, 2009). tRNAs foram localizados usando tRNAscan-SE e os RNA ribossomais foram localizados usando BLASTn.

H. seropedicae tem o potencial genético de prover benefícios ao se associar com plantas, estimulando o crescimento vegetal tanto como consequência da fixação de nitrogênio quanto pela produção de fitohormônios como o etileno (PEDROSA *et al*., 2011).

É capaz de crescer em monossacarídeos como D-glicose, D-frutose, D-galactose e L-arabinose, com álcoois de açúcar e acidos orgânicos como L-malato ou L-lactato. O genoma contem o conjunto completo de genes para as vias Entner-Doudoroff e para fosfato pentose. A via Embden-Meyerhoff-Parnas (EMP) não possui o gene que codicifica para 6-fosfofrutoquinase, sugerindo que provavelmente requer o envolvimento das vias Entner-Doudoroff e para fosfato pentose para metabolizar D-frutose ou D-manose para piruvato (PEDROSA *et al*., 2011).

Os genes para a fixação de Nitrogênio de *H. seropedicae*, incluindo *nif*A, *nif*B, *nif*Z, *nif*Z1, *nif*H, *nif*D, *nif*K, *nif*E, *nif*N, *nif*X, *nif*Q, *nif*W, *nif*V, *nif*U e *nif*S foram encontrados, intercalados com os genes *fix*, *mod*, *hes*, *fdx* e *hsc*. Esse cluster gênico contem todos os genes necessários para a síntese e atividade da nitrogenase (PEDROSA *et al*., 2011).

H. seropedicae é capaz de colonizar várias *Gramineae* (RONCATO-MACCARI, 2003). Os genes que codificam enzimas de degradação da parede celular de plantas, tais como glicosidases, celulases e hemicelulases associadas à penetração bacteriana não foram encontrados no genoma de *H. seropedicae*. É provável, portanto, que este organismo dependa apenas de descontinuidades naturais da raiz da planta epiderme para penetração (OLIVARES *et al*., 1997).

1.2. *Azospirillum brasilense*

A. brasilense são bactérias fixadoras de nitrogênio, aerobicas, Alfaproteobactérias, Gram-negativas, com flagelo polar em meio líquido e cílios laterais em meios sólidos. Elas são capazes de realizar todas as vias do ciclo do nitrogênio, exceto a nitrificação. São quimio-organotróficas típicas, podem colonizar os tecidos de milho, trigo, arroz, sorgo e plantas não gramíneas como endofíticos facultativos (DÖBEREINER, PEDROSA, 1987).

Azospirillum é uma bactéria que promove o crescimento vegetal de diferentes maneiras, incluindo a Fixação Biológica de Nitrogênio, deaminação do 1-aminociclopropano-1-carboxilato, produção de óxido nítrico e produção de fitohormônios (BASHAN *et al*., 2004; BLAHA *et al*., 2006; SPAEPEN *et al*., 2007).

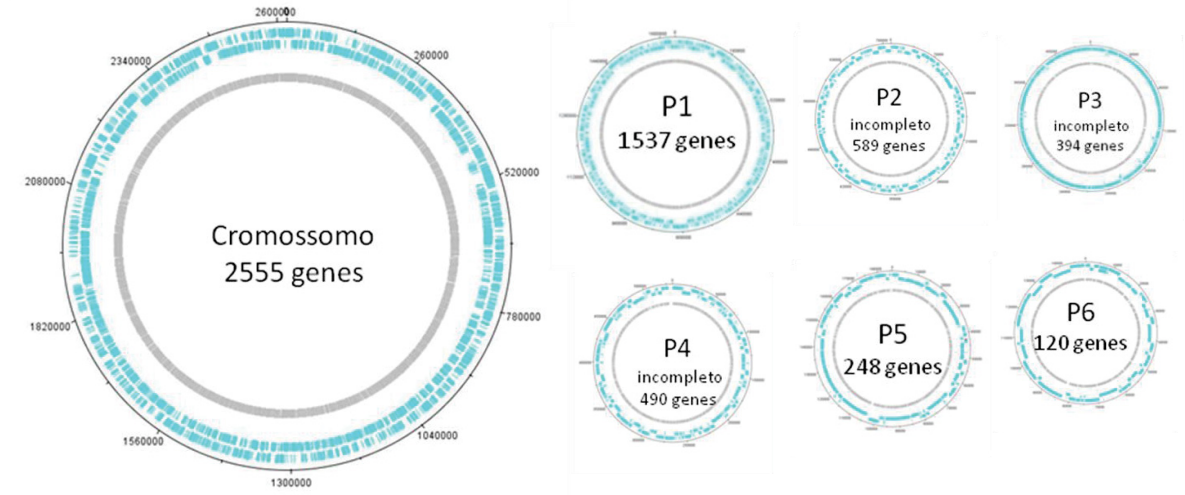


FIGURA 01: REPRESENTAÇÃO DOS SETE REPLICONS DE *A. brasilense* FP2.

O genoma do *A. brasilense* FP2 ainda não está publicado, mas já foi sequenciado, montado parcialmente e anotado pelo grupo GENOPAR. A versão de montagem do genoma realizada por Dieval Guizelini consiste de sete replicons: 1 cromossomo, com 2555 genes anotados e seis plasmídeos. O plasmídeo P1 com 1537 genes anotados, o plasmídeo P2 com 589 genes anotados, o plasmídeo P3 com 394 genes anotados, o plasmídeo P4 com 490 genes anotados, o plasmídeo P5 com 248

genes anotados e o plasmídeo P6 com 120 genes anotados (GUIZELINI, 2016), sendo um total de 5933 genes, conforme a Figura 01.

1.3. Fixação Biológica de Nitrogênio

O nitrogênio é essencial para todos os seres vivos, pois é integrante das biomoléculas. Está presente no ar atmosférico em grande quantidade (80%) na forma N_2, porém dessa forma não pode ser assimilado pela maioria dos seres vivos, devido à presença de uma forte ligação tríplice entre as duas moléculas de N. Sendo assim esses organismos o absorvem na forma de outras substâncias que possuem íons amônio.

Microrganismos diazotróficos realizam a fixação biológica de nitrogênio, isto é, eles são capazes de reduzir o nitrogênio atmosférico à amônio. Para que isso ocorra, o microrganismo deve estar diante de condições favoráveis tais como, baixas concentrações de nitrogênio fixado, baixas tensões de oxigênio e temperatura adequada (POSTGATE, 1982).

A fixação de nitrogênio consiste na reação de redução do N_2 a NH_3, catalisada pelo complexo enzimático nitrogenase cuja estequiometria é mostrada abaixo (SIMPSON & BURRIS, 1984).

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

A regulação da fixação de nitrogênio pode ocorrer através da regulação transcricional da nitrogenase e da regulação pós traducional. Esses dois tipos de regulação respondem aos níveis de amônio e oxigênio, e a proteína NifA e as proteínas do sistema Ntr estão diretamente envolvidas com essa regulação.

1.3. Controle do metabolismo geral de Nitrogênio pelo sistema Ntr

O sistema Ntr controla o metabolismo geral de nitrogênio, regulando a utilização de amônio e de fontes alternativas de nitrogênio como nitrato, aminoácidos e N_2 em proteobactérias (DRUMMOND *et al.,*1983).

Em enterobactérias o sistema Ntr é composto por sete proteínas: NtrB (produto do gene *ntrB*), NtrC (produto do gene *ntrC*), GlnD ou UTase (produto do gene *glnD*), GlnB (produto do gene *glnB*), GlnE ou ATase (produto do gene *glnE*), GlnK (produto do gene *glnK*) e glutamina sintetase (GS) ou GlnA (produto do gene *glnA*). A proteína AmtB (produto do gene *amtB*) também tem sido considerada como integrante do sistema Ntr de *E. coli* (JAVELLE & MERRICK, 2005).

A enzima GS, que catalisa a síntese de glutamina a partir de glutamato e amônio, é regulada por adenilação/desadenilação, catalisada pela enzima ATase, codificada pelo gene *glnE* (ARCONDÉGUY, 2001). A enzima ATase, por sua vez, tem sua atividade regulada pela proteína PII, que é regulada por uridililação/desuridililação, que é catalisada pela proteína GlnD (ADLER *et al.,* 1975). Quando a célula se encontra em condições limitantes de nitrogênio, PII e NtrC encontram-se em altas concentrações. Nestas condições, GlnD modifica covalentemente PII por adição de um grupo uridilil em um resíduo específico de tirosina em cada uma das subunidades da proteína. A proteína PII uridililada resultante promove então a desadenilação de GS via ATase. PII uridililada não é capaz de interagir com NtrB, o que culmina com a fosforilação de NtrC (REITZER, 2003). Nesse processo, NtrB é inicialmente autofosforilada no resíduo His-139, numa reação que envolve hidrólise de ATP (NINFA & MAGASANIK, 1986). Em ambiente rico em nitrogênio, GlnD remove o grupamento uridilil de PII, que promove adenilação de GS via ATase. Nestas condições, a expressão de *glnA* se encontra em um nível basal. PII, no estado não uridililada, interage com NtrB, que catalisa então a desfosforilação de NtrC. NtrC desfosforilada não é capaz de ativar a transcrição (REITZER, 2003).

O sistema Ntr também é regulado por α-cetoglutarato, que se constitui no principal sinal de carbono em *E. coli*. Essa molécula regula a atividade de PII ligando-se diretamente a ela (KAMBEROV *et al.,* 1995). Cada trímero de PII possui três sítios para ligação de alfa-cetoglutarato e três para a ligação de ATP. A ligação de alfa-

cetoglutarato e ATP na proteína PII é cooperativa (ARCONDÉGUY *et al.,* 2001). Em concentrações fisiológicas de alfa-cetoglutarato, três moléculas de ATP ligam-se firmemente a PII (KAMBEROV *et al.,* 1995). Em contraste, quando ATP está em concentrações fisiológicas, a primeira molécula a se ligar na proteína PII é o α-cetoglutarato. Essa ligação exerce cooperatividade negativa na ligação de moléculas de alfa-cetoglutarato adicionais na proteína PII (KAMBEROV *et al.,* 1995). A ligação de PII regula a habilidade de PII de interagir com ATase. A forma completamente saturada de ATP e alfa-cetoglutarato não é capaz de interagir com qualquer outra proteína (KAMBEROV *et al.,* 1995). Dessa forma, o sinal de carbono pode ir contra o sinal de nitrogênio, resultando na integração de dois tipos de informação. (NINFA e ATKINSON, 2000).

A NtrC é uma proteína dimérica, na qual cada subunidade possui massa em torno de 55 kDa (MERRICK & EDWARDS, 1995). NtrC faz parte do grupo de proteínas ativadoras de transcrição dependentes de σ^{54} (KUSTU *et al.,* 1989; MORETT e SEGOVIA, 1993). Como é característico nesse grupo de proteínas ativadoras de transcrição, NtrC possui três domínios distintos: o N-terminal, com uma atividade regulatória, o central, que é o domínio catalítico, interage diretamente com a subunidade σ^{54} da RNA polimerase e possui um sítio de ligação de ATP, e o C-terminal, que possui um motivo hélice-volta-hélice de ligação ao DNA, e por isso acredita-se que este domínio seja o responsável pela interação da região promotora do gene a ter sua transcrição ativada e a proteína (DRUMMOND *et al.,* 1986). Apenas NtrC fosforilada é capaz de ativar a transcrição a partir de promotores alvo. A fosforilação parece não estar relacionada com a ligação de NtrC ao DNA, mas sim a oligomerização dos dímeros de NtrC no sítio de ligação desta proteína no DNA (REITZER, 2003). Os genes que tem sua transcrição ativada por NtrC incluem *glnA, ntrBC*, *glnHPQ* (envolvidos no transporte da glutamina), *argT* (envolvido no transporte de arginina), *hisJQMP* (envolvidos no transporte de histidina), *nasFEDCBA* (envolvidos na assimilação de nitrato) e o gene *nifA*, cujo produto regula a transcrição dos genes *nif* (MERRICK e EDWARDS, 1995).

As proteínas do tipo PII, GlnB e GlnZ (*A. brasilense*) e GlnB e GlnK (*H. seropedicae*) têm um papel fundamental na coordenação das vias do metabolismo de nitrogênio. O mutante *glnZ* de *A. brasilense* é capaz de fixar nitrogênio como observado na estirpe selvagem (DE ZAMAROCZY, 1998), já o mutante *glnB* e incapaz de fixar

nitrogênio. O duplo mutante *glnBglnZ* apresenta ainda deficiência de crescimento em meio rico ou mínimo, este fenótipo é complementado com o gene *glnB* ou com o gene *glnZ*, indicando que a presença de uma das proteínas PII é necessária para o crescimento celular (DE ZAMAROCZY, 1998). Mutantes *glnZ* tem uma taxa de transporte de amônio maior que o selvagem (DE ZAMAROCZY, 1998). O mutante *glnK* de *H. seropedicae* não fixa nitrogênio provavelmente devido a falta de atividade da proteína NifA neste mutante. Já a mutação do gene *glnB* não altera a fixação de nitrogênio neste organismo (NOINDORF, 2006).

1.4. Controle da expressão do complexo enzimático nitrogenase

A proteína ativadora de transcrição dos genes *nif*, que atua em conjunto com a RNA polimerase - σ54, é a NifA. Esse mecanismo de ativação foi encontrado em todos os diazotrofos pertencentes à classe das proteobactérias estudados até o momento (DIXON e KAHN, 2004). Sendo uma proteína da família das EBP, a proteína NifA liga-se a uma seqüência conservada no DNA, denominada UAS ("Upstream Activator Sequence") (5'-TGT-N10-ACA-3'). Esta seqüência está localizada a uma distância de 80 a 150 nucleotídeos a montante do sítio de início da transcrição (MORETT *et al.,* 1988; MORETT e BUCK, 1988). Souza e colaboradores identificaram e seqüenciaram o gene *nifA* de *Herbaspirillum seropedicae* (SOUZA *et al.,* 1991). A proteína NifA das proteobactérias diazotróficas sofre regulação pós-traducional em resposta a dois sinais ambientais: os níveis de oxigênio e nitrogênio fixado.

1.5. Controle da atividade do complexo enzimático nitrogenase

A adição de amônio a células fixando nitrogênio causa um rápido desligamento da atividade da nitrogenase. Em alguns diazotrofos como *A. brasilense* o controle pós-traducional da nitrogenase, envolve a modificação covalente da proteína Fe ou dinitrogenase redutase (ZHANG *et al*., 1997). Quando amônio é adicionado a células de *A. brasilense* uma das subunidades da proteína Fe é ADP-ribosilada pela dinitrogenase redutase ADP-ribosil transferase (DraT, produto do gene *draT*) (POPE *et al*., 1985;

LOWERY *et al*., 1986; LOWERY e LUDDEN, 1988; JOUANNEAU *et al*., 1989). Quando o amônio é consumido, a proteína dinitrogenase redutase glicohidrolase (DraG, produto do gene *draG*) remove o grupamento ADP-ribosil da dinitrogenase redutase, restaurando sua atividade (LUDDEN e BURRIS, 1976; SAARI et al., 1984; SAARI *et al.,* 1986). Em *A. brasilense* a proteína AmtB é necessária para que ocorra modificação da dinitrogenase em resposta a adição de amônio (KLASSEN *et al*., 2001; KLASSEN *et al*., 2005; HUERGO *et al*., 2006). Mutantes cromossomais *glnB* e *glnZ* de *A. brasilense* apresentaram uma modificação pós-traducional da proteína NifH, em resposta a amônio, diferente da observada na estirpe selvagem. (HUERGO *et al*., 2006). A proteína GlnB é necessária para a ADP-ribosilação da proteínas NifH (inativação de NifH) na presença de amônio, sugerindo que GlnB e necessária para a ativação de DraT. GlnZ e necessária para a reativação de NifH (retirada do ADP-ribosil) (HUERGO *et al*., 2006).

A *H. seropedicae* não possui o sistema DraT/DraG e o mecanismo de regulação pós-traducional da nitrogenase permanece desconhecido. Noindorf e colaboradores (2006) mostraram que a proteína AmtB é essencial para o desligamento completo da nitrogenase por adição de íons amônio em *H. seropedicae.*

2. Objetivo Geral

Este trabalho possui dois objetivos gerais:

1. Determinar o transcriptoma de *A. brasilense* sob condições de limitação de nitrogênio;

2. Finalizar a Re-Anotação do genoma de *H. seropedicae* SmR1 utilizando dados de RNA-Seq.

2.1. Objetivos Específicos

1. Identificar genes expressos de *A. brasilense* nas diferentes condições utilizando ferramentas de Bioinformática.

1.1: Filtrar o arquivo de sequências para remoção de sequências de adaptador e de baixa qualidade;

1.2: Mapear as sequências filtradas no genoma de referência da *A. brasilense*, utilizando o pacote de programas CLC Genome WorkBench;

1.3: Quantificar os níveis de expressão para cada gene;

1.4: Comparar os níveis de expressão gênica entre as diferentes condições testadas e obter uma lista com os genes diferencialmente expressos.

2. Construir uma rede metabólica de *A. brasilense,* utilizando os dados obtidos.

2.1: Identificar as enzimas encontradas utilizando o banco de dados KEGG;

2.2: Reconstruir as vias metabólicas a partir das enzimas identificadas.

3. Reanotar o genoma da *H. seropedicae* SmR1 utilizando dados de RNA-Seq.

3.1: Identificar novos genes;

3.2: Identificar operons;

3.3: Atualizar a classificação funcional dos genes.

3. Materias e Métodos

Os dados de RNA-Seq foram obtidos utilizando o Sistema SOLiD (Sequencing by Ligation and Detection) que usa DNA ligase para sequenciamento. O SOLiD utiliza octâmeros marcados com fluoróforos para identificar a sequência alvo. As primeiras 5 bases da sonda garantem a especificidade da ligação com o *template*, enquanto que as 3 útimas bases são inosinas que anelam de maneira inespecífica. Conectado à última inosina temos o fluróforo que gerará o sinal luminoso a ser detectado pelo sequenciador. Cada fragmento é amplificado milhares de vezes na superfície de uma *bead*. Essas *beads* são fixadas em uma lâmina de vidro e são identificadas através das coordenadas do ponto luminoso na lâmina (VARUZZA, 2013).

O sequenciamento realizado pela tecnologia SOLiD gera dois arquivos. O arquivo com extensão '.csfasta' (*color space*) e o arquivo com extensão '.qual'. No arquivo '.qual' estão os valores de qualidade associados a cada base.

Os dados obtidos foram mapeados ao genoma de *A. brasilense* FP2 utilizando o *software* CLC Workbench (CLC, 2014) e analisados. Os dados do sequenciamento passaram por um filtro de qualidade e foram mapeados em relação ao genoma de *A. brasilense* FP2 usando o programa CLC Workbench e as condições de cultivos foram comparados entre si. Os genes diferencialmente expressos foram selecionados levando em consideração o *fold-change* maior ou igual a dois e o *e-value* menor que 0,05 (RAMOS *et. al*, 2011).

As sequências obtidas foram mapeadas em relação ao genoma da espécie e a quantificação de cada gene expresso foi normalizada levando em consideração o tamanho do gene, o número de leituras do gene e da biblioteca (total de sequências mapeadas) que é o método de RPKM (*Reads Per Kilobase per Million mapped reads*, $RPKM=(C*10^{9})/(L*N)$ em que C é o numero de leituras por gene, L é o tamanho do gene e N é o tamanho da biblioteca) (MORTAZAVI *et al*., 2008). A cobertura calcula quantas vezes a quantidade de leituras pode cobrir o tamanho do gene. Cobertura = (C*N)/L, onde C = número de leituras por gene, L = tamanho do gene e N é o tamanho das leituras.

Para visualizar o empilhamento das leituras em relação ao genoma no *software* Artemis (SANGER, 2012) foi feita a conversão dos arquivos para o formato 'bam' usando as ferramentas contidas no pacote Samtools (SAMTOOLS, 2012).

O transcriptoma foi analisado utilizando o pacote de *software* DESeq (ANDERS & HUBER, 2010) e o EdgeR (RPBINSON & OSHLACK, 2010). Foi utilizado o KEGG Pathway

(KEGG, 2014) para identificar em quais vias metabólicas os genes diferencialmente expressos se encontram.

O fluxograma da etapa do projeto de análise de expressão diferencial é mostrado na Figura 02.

FIGURA 02: FLUXOGRAMA DAS ETAPAS DO PROJETO.

3.1. Descrição das Condições de Cultivo da *H. seropedicae* SmR1

O Quadro 01 a seguir, descreve as condições de cultivo das bibliotecas de RNA-Seq utilizadas para verificar a expressão de genes. Para cada condição de cultivo foi utilizada duas replicatas. Os dados destas bibliotecas foram utilizados para a etapa de re-anotaçao do genoma de *H. seropedicae* SmR1.

QUADRO 01: DESCRITIVO DAS CONDIÇÕES DE CULTIVO DAS BIBLIOTECAS DE RNA-SEQ OBTIDOS COM A TECNOLOGIA SOLID.

Condições	Descrição	Referência
1. NFbHPN-Malato	A condição controle consiste nas bactérias cultivadas a 30ºC em meio NFbHPN, utilizando malato como fonte de carbono até D.O 0,8.	(TADRA-SFEIR, 2011)
2. NFbHPN-Malato-Naringenina	Bactérias cultivadas a 30ºC em meio NFbHPN, utilizando malato como fonte de carbono na presença de 100µM do flavonóide naringenina, até D.O. 0,8.	(TADRA-SFEIR, 2011)
3. NFbHP-Malato-Nitrato	Bactérias cultivadas a 30ºC em meio NFbHP, utilizando malato como fonte de Carbono e nitrato 10mM como fonte de nitrogênio.	(BONATO, 2012)
4. Planctônico–milho-1dia	Bactérias recuperadas do meio líquido do sistema hidropônico (denominadas plantônicas) após um dia de inoculação em raiz de milho.	(BALSANELLI, 2013)
5. Planctônico-milho-3dias	Bactérias recuperadas do meio líquido do sistema hidropônico (denominadas plantônicas) após um dia de inoculação em raiz de milho.	(BALSANELLI, 2013)
6. Aderido–milho- 1dia	Bactérias recuperadas da superfície radicular de milho (denominadas aderidas) após um dia de inoculação.	(BALSANELLI, 2013)
7. Aderido-milho3dias	Bactérias recuperadas da superfície radicular de milho (denominadas aderidas) após três dias de inoculação.	(BALSANELLI, 2013)
8. Planctônico- trigo	Bactérias recuperadas do meio líquido	(PANKIEVICZ, 2013)

Condições	Descrição	Referência
	do sistema hidropônico (denominadas plantônicas) após três dias de inoculação em raiz de trigo.	
9. Aderido-trigo	Bactérias recuperadas da superfície radicular de trigo (denominadas aderidas) após três dias de inoculação.	(PANKIEVICZ, 2013)
10. NFbHPN-Malato-Alto Oxigênio	Bactérias cultivadas a 30ºC em meio NFbHPN, utilizando malato como fonte de carbono até D.O. de 0,4. Depois de atingida a D.O. 0,4 as bactérias continuaram sendo cultivadas nesta condição por mais 1,5 horas.	(BATISTA, 2013)
11. NFbHPN-Malato-Baixo Oxigênio	Bactérias cultivadas a 30ºC em meio NFbHPN, utilizando malato como fonte de carbono até D.O. de 0,4. Nesse momento as bactérias passaram a ser cultivadas em condições microanaerobicas (concentração inicial de oxigênio 2%) por 1,5 horas.	(BATISTA, 2013)

3.1. Descrição das Condições de cultivo para análise de expressão diferencial em condição limitante de amônio de *Azospirillum brasilense* FP2

A. brasilense FP2 foi cultivada em duas condições: condição de fixação e de não fixação. As descrições das condições foram fornecidas pela Dra. Michelle Zibetti Tadra Sfeir.

Para condição de fixação: as bactérias foram crescidas durante 16 horas em 50 mL de meio NFbHPN (10 mmol/L de NH4Cl) lactato, sob agitação constante e 30°C. As células foram centrifugadas a 4.000 rpm por 5 minutos, à temperatura ambiente, e ressuspensas em 20 mL de meio **NFbHP** lactato com 1% de LB, para que atingisse uma Densidade Óptica 600 nm de 1.2.

Estes 20 mL foram então transferidos para um frasco de 65 mL de volume total e então incubados à 30°C sob agitação a 120 rpm, por 4 horas. Após este período foi injetado 4,5 mL de acetileno, incubados novamente a 30°C e 120 rpm por 20 minutos e posteriormente feito detecção da atividade da nitrogenase através da análise da produção de etileno.

Para Condição de não fixação: as bactérias foram crescidas durante 16 horas em 50 mL de meio NFbHPN (10 mmol/L de NH4Cl) lactato, sob agitação constante e 30°C. As células foram centrifugadas a 4.000 rpm por 5 minutos, à temperatura ambiente e ressuspensas em 20 mL de meio **NFbHPN** lactato com 1% de LB, para que atingisse uma Densidade Óptica 600 nm de 1.2.

Estes 20 mL foram então transferidos para um frasco de 65 mL de volume total e então incubados a 30°C sob agitação a 120 rpm, por 4 horas. Após este período foi injetado 4,5 mL de acetileno, incubados novamente a 30°C e 120 rpm por 20 minutos e posteriormente feito detecção da atividade da nitrogenase através da análise da produção de etileno.

A.brasilense FP2 teve seu RNA total coletado e sequenciado pela plataforma Solid, nas condições de fixação e de não fixação, em replicatas.

3.2. *Softwares*

Os *softwares* descritos a seguir foram utilizados para as etapas de remontagem do genoma, análise comparativa do transcriptoma da *A. brasilense* FP2 e reanotação do genoma da *H. seropedicae* SmR1.

3.2.1 DOOR 2.0 (a Database for prOkaryotic OpeRons – Banco de Dados para Operons em Procariotos.)

O DOOR (Database for prOkaryotic OpeRons) é um banco de dados de operons procarióticos atualizado regularmente (http://csbl.bmb.uga.edu/DOOR), que está integrado com um preditor de operons (MAO et al., 2009; DAM et al., 2007). Apresenta os operons e suas funções através de um visualizador interativo (MAO et al., 2013).

O programa de predição de operons emprega dois classificadores, um para genomas com operons validados experimentalmente e um para genomas com apenas dados limitados. Para o

primeiro caso, o programa foi treinado em um subconjunto de operons conhecidos; enquanto que no segundo caso, o programa foi treinado com base apenas nas características gerais dos genomas.

3.2.2 Artemis 16.0

O Artemis é um navegador interativo de genomas, gratuito, disponibilizado pelo Instituto Sanger (http://www.sanger.ac.uk/science/tools/artemis). Facilita a anotação de genomas e integrado a ferramenta 'bamview' (CARVER *et al.* 2010; CARVER *et al.* 2012) permite visualizar dados de sequenciamento feitos em plataformas de nova geração. Pode ler arquivos nos formatos EMBL, GenBank ou sequência em FASTA, FASTA indexados, GFF ou formato cru (RUTHERFORD, 2000). Deve funcionar em qualquer sistema com uma versão recente do Java.

3.2.3 CLC DNA Workbench 7.5.5

O CLC DNA *Workbench* é uma bancada de trabalho 'virtual', desenvolvido pela QIAGEN (https://www.qiagenbioinformatics.com/products/clc-genomics-workbench), que possibilita análises em sequencias de DNA, RNA e proteínas. O *software* suporta ler mapeamento de Sanger e dados de sequenciamento das tecnologias SOLiD, Ion Torrent, 454 e Illumina; montagem de novo, detecção de SNP em genomas inteiros, a detecção de variações estruturais dos genomas, análise de Chip-seq, RNA-Seq, análise de *small*RNA e BLAST. Permite analisar, visualizar e comparar dados genômicos, transcriptomas e epigenômicos de todas as principais plataformas de *high-throughput* (CLC BIO, 2012).

3.2.4 Samtools 0.1.19

SamTools (LI, 2009) fornece vários utilitários para manipulação de alinhamentos no formato SAM (Sequence Alignment/Map), incluindo separar, mesclar, indexar, converter em

outros formatos e gerar alinhamentos em um formato por posição. Disponível em (http://samtools.sourceforge.net/).

O arquivo de mapeamento do programa CLC DNA Workbench 7 foi exportado no formato ".sam" e foi convertido para o modo binário (extensão .bam) com a seguinte sequência de comandos:

1) samtools view –bS arquivo.SAM > arquivo.BAM (Conversão do mapeamento do formato SAM para o formato BAM);

2) samtools sort arquivo.BAM arquivobam.sorted (Alinhamento das coordenadas do mapeamento) ;

3) samtools index arquivo.sorted.bam (Classifica o índice para que o acesso ao arquivo seja mais rápido).

3.2.5 SHRiMP 2.2.3 (SHort Read Mapping Package)

SHRiMP é um pacote de *softwares* para alinhamento de *reads* contra o genoma alvo. Disponível em (http://compbio.cs.toronto.edu/shrimp/). Ele foi desenvolvido primeiramente tendo em mente os numerosos *reads* de máquinas de sequenciamento de próxima geração. O método baseia-se em uma técnica de mapeamento rápido dos *reads* (RUMBLE *et al*, 2009).

As sequências foram mapeadas no genoma da bactéria *A. brasilense* FP2 utilizando a ferramenta SHRiMP, com o seguinte comando de execução:

gmapper-cs arquivo.csfasta referencia.fasta –N8 –o1 –h80% -E 2>arquivo.log>arquivo.sam

Neste comando o parâmetro –N se refere ao número de núcleos de processamento que foi utilizado pelo computador, sendo então, 8 núcleos. O parâmentro -o determina o número de vezes que um read pode ser alinhado, neste caso, apenas 1 vez. O parâmetro -h estipula a porcentagem de homologia entre as sequências, sendo de 80%. E –E permite o arquivo de saída no formato sam.

3.2.6 KEGG e KAAS

O KEGG (*Kyoto Encyclopedia of Genes and Genomes* - Kyoto Enciclopédia de Genes e Genomas) é um banco de dados para a compreensão das funções de sistemas biológicos, a partir de informações em nível molecular, especialmente em conjuntos de dados moleculares gerados pelo sequenciamento do genoma e outras tecnologias experimentais de elevado rendimento (KANEHISA *et al.*, 2013; KANEHISA *et al.*, 2012). Desenvolvido pelo laboratório Kanehisa e disponível em (http://www.kegg.jp/).

O KAAS (KEGG *Automatic Annotation Server* - KEGG Servidor de anotação automática), (http://www.genome.jp/tools/kaas/) fornece anotação funcional de genes por meio de comparações BLAST contra o banco de dados manualmente curado GENES KEGG. O resultado contém KO (KEGG *Orthology*) que é um identificador automatizado de atribuições e vias metabólicas (MORIYA *et al.,* 2007).

3.2.8 Quality Assessment 0.5

Com o *software* Quality Assessment pode-se obter e analisar gráficos que mostram a distribuição dos valores de qualidade a partir do sequenciamento. Recebe dois arquivos de entrada: um com os valores de qualidade e outro com as sequencias de nucleotídeos ou color space (SOLiD). E gera dois novos arquivos de saída, com os dados dentro dos parâmetros de qualidade requeridos.

Este *software* permite adotar as normas de qualidade mais rigorosas para os dados de sequência, com base na análise de qualidade de gráficos e cobertura estimada após a aplicação do filtro de qualidade, proporcionando uma sequência aceitável para as inferências biológicas. Aplicação de filtros de qualidade utilizando o *software* Quality Assessment, proporciona uma maior precisão na definição dos parâmetros de corte e no transcritoma como um todo. Disponível em: https://sourceforge.net/projects/qualevaluato (RAMOS *et al*, 2011).

3.2.9 FastQC 0.11.5

Desenvolvido pelo Instituto Babraham de Bioinformática em linguagem Java. Antes de analisar dados de sequenciadores de nova geração para tirar conclusões biológicas é necessário realizar alguns testes de controle de qualidade para garantir que os dados brutos sejam confiáveis

e que não há problemas nos dados que podem afetar a interpretação. O FastQC realizada análises de qualidade na forma de gráficos. Realizada a estatísticas básica dos dados, a qualidade da sequencia por base, a distribuição por tamanho de sequências dentre outras análises. Está disponível em: www.bioinformatics.babraham.ac.uk/projects/fastqc/

3.2.10 Rockhopper 2.0.3

O Rockhopper foi desenvolvido em Java, sendo um sistema abrangente e de interface amigável para a análise computacional de dados de RNA-seq bacteriano. Rockhopper faz *upload* de arquivo(s) de reads de sequenciamento em larga escala nos formatos Fastq, Qseq, Fasta, Sam, ou Bam.

Rockhopper realiza as seguintes tarefas: alinhamento de *reads* a um genoma, normalização de dados de diferentes experimentos, montagem de transcriptoma, identifica os limites de transcrição e novas transcrições, como RNAs curtos, quantifica expressão, expressão diferencial de genes, localização de operons e visualização dos resultados em um navegador de genoma (MCCLURE *et al.*, 2013). E está disponível no endereço http://cs.wellesley.edu/~btjaden/Rockhopper/ .

No programa são utilizadas duas funcionalidades para estimar a probabilidade de genes serem operons: a distância entre os genes e a similaridade de expressão dos genes nos dados de RNA-seq. Os recursos são combinados, junto com a probabilidade de ser um operon, usando um classificador Naive Bayes com operons prévios para determinar a probabilidade de serem co-transcritos. O desempenho do sistema foi avaliado com os dados obtidos de RNA-Seq realizados com cinco bactérias diferentes (MCCLURE *et al.*, 2013).

O programa foi utilizado com suas configurações e parâmetros padrões para prever operons com base em dados de RNA-Seq em 11 condições de cultivo da bactéria *H. seropedicae*.

3.2.11. DESeq 1.32.0

O DESeq é um pacote para análise de Expressão Diferencial da Bioconductor programado em linguagem R (pode ser realizado *download* em: https://bioconductor.org/packages/release/bioc/html/DESeq.html).

O DESeq estima o tamanho efetivo da biblioteca ou normalização. A informação de tamanho de biblioteca é chamada de vetor de fatores de tamanho, uma vez que o pacote só precisa conhecer os tamanhos de biblioteca relativos. Se as contagens de genes numa amostra são, em média, duas vezes mais altas do que na outra (porque a biblioteca foi sequenciada duas vezes), o fator de tamanho para a primeira amostra deve ser o dobro da outra amostra. A inferência no DESeq se baseia numa estimativa da relação típica entre a variância dos dados e sua média (ANDERS, S. & HUBER, 2010; ANDERS, S.; REYES, A. e HUBER, 2012).

3.2.12. EdgeR 3.22.3

O EdgeR é um pacote para análise de Expressão Diferencial da Bioconductor programado em linguagem R (pode ser realizado *download* em: http://bioconductor.org/packages/release/bioc/html/edgeR.html).

EdgeR trabalha na análise de expressão diferencial e não na quantificação da expressão. Por exemplo, pode-se esperar que as contagens de leitura sejam proporcionais ao comprimento do gene, mas o EdgeR não necessita geralmente ajustar o comprimento do gene porque o comprimento do gene tem a mesma influência relativa nas contagens lidas para cada amostra de RNA. EdgeR normaliza a composição de RNA ao encontrar um conjunto de fatores de escala para os tamanhos de biblioteca que minimizam as alterações em log-fold entre as amostras para a maioria dos genes. O método padrão para calcular esses fatores de escala usa uma média aparada de M-values (trimmed mean of M-values - TMM) entre cada par de amostras (ROBINSON & OSHLACK, 2010).

3.2.13. Baggerley's test

O Baggerley's test está integrado ao programa CLC DNA Workbench 7.5.5 e compara as proporções de contagens em um grupo de amostras com as de outro grupo de amostras e é adequado para casos em que há réplicas disponíveis nos grupos. As amostras recebem pesos diferentes dependendo do seu tamanho (contagens totais). Os pesos são obtidos assumindo uma distribuição Beta sobre as proporções em um grupo e estimando-as, juntamente com a proporção

de uma distribuição binomial. O resultado é uma estatística de teste do tipo t ponderada (BAGGERLY *et al.,* 2003).

3.2.14. MUMmer 3.0

MUMmer é um sistema para alinhar genomas, seja em forma completa ou em forma de *draft*. Diponível em (http://mummer.sourceforge.net/), (KURTZ, 2004). Pode ser utilizado para encontrar inversões, duplicações e deleções em comparações de genomas. Podendo assim inferir o quão os genomas são semelhantes.

3.2.15. HGF

Realiza anotação do genoma a partir do formato fasta, o programa HGF (Hybrid Gene Finder – Localizador Hibrido de Gene) está disponível no *site* http://www.bioinfo.ufpr.br/hgf/, (RAITTZ, 2010). Utiliza técnicas de Inteligencia Artificial para Reconhecimento de Padrão, através de um Algoritmo Genético que obtém os dados e a classificação é feita por uma Rede Neural Artificial, cujo protótipo foi construído no Matlab®.

3.2.16. SILA

É um sistema para anotação automática de genomas de procariotos que engloba as estratégias de GeneFinding e de comparação de sequências com o RAFTS3 (Rapid Alignment Free Tool for Sequences Similarity Search), (VIALLE, 2016). Os genes preditos são comparados com bancos de dados de sequências de proteínas (NR, Pfam e COG) utilizando o RAFTS3 para buscar informações que possam predizer as funções dos genes, através da semelhança de sequências. O SILA apresenta resultados de anotação com altas taxas de acerto na predição e identificação dos genes (VIALLE, 2013). O serviço de anotação está disponível em http://200.236.3.34/SILA/login.jsp.

3.2.17. BLAST

O BLAST (Basic Local Alignment Search Tool) realiza busca de regiões de similaridade local entre sequências (ALTSCHUL *et al.,* 1990). O programa compara sequências de nucleotídeos (BLASTn) ou proteínas (BLASTp), ou traduz a sequencia de nucleodídeos e busca por proteínas (BLASTx), em bancos de dados de sequências e calcula a significância estatística dos alinhamentos. O BLAST pode ser usado para inferir relações funcionais e evolutivas entre sequências, bem como ajudar a identificar membros de famílias de genes. Também está integrado ao BLAST a busca por domínios preteícos conservados.

3.2.18. FGAP

O FGAP (Fecha Gap) é uma ferramenta para fechar *gaps* de sequências genômicas, que aproveita diferentes conjuntos de dados, combina novos dados de sequenciamento com *draft* de genoma. O FGAP usa BLAST para alinhar os *contigs* contra um conjunto de *draft* do genoma, com o objetivo de encontrar sequências que se sobrepõem aos *gaps*. O algoritmo seleciona a melhor sequência para preencher e eliminar o *gap* (PIRO *et al.,* 2014).

3.2.19. G-FINISHER

Padrões biólogos observados em sequências genômicas permite identificar possíveis regiões montadas equivocadamente; sendo possível reorganizar as sequências com a intenção de melhorar a montagem do genoma. No G-FINISHER os *contigs* são quebrados nos pontos de máximo e mínimo local da curva Fuzzy-GC-Skew e armazenados em nós de um grafo sem bordas (GUIZELINI, 2016). Esses nós são ordenados com base em uma sequência de referência e submetidos para fechamento de *gaps* pelo jFGap. O *software*, escrito em Java, está disponível em http://gfinisher.sourceforge.net/.

3.2.19. GLIMMER 3.02

O *software* GLIMMER (Gene Locator and Interpolated Markov ModelER) localiza genes em genomas microbianos utilizando Modelos de Markov Interpolados (SALZBERG *et al*., 1998; DELCHER *et al*., 1999). Está disponível em (https://ccb.jhu.edu/software/glimmer/).

4. Resultados e Discussões

4.1 Transcriptoma do *Azospirillum brasilense* FP2 crescido em condições de fixação de Nitrogênio

A *Azospirillum brasilense* é uma bactéria utilizada como biofertilizante, e o nitrogênio fixado por esta batéria pode ser utilizado pela planta hospedeira. O estudo do transcriptoma de *A. brasilense,* crescido em condições de fixação de nitrogênio, nos permite determinar qual o conjunto de genes que tem a sua expressão regulada quando a bactéria é crescida nessas condições.

O *A. brasilense* FP2 foi crescido em condições de fixação de Nitrogênio e a expressão e atividade do complexo enzimático da nitrogenase foi confirmada pela determinação da produção de etileno. O transcriptoma do FP2 crescido nessa condição foi obtido, assim como o transcriptoma do FP2 crescido em condições de alto nitrogênio. Inicialmente mapeamos os dados de transcriptoma contra a sequência do genoma de *A. brasilense* SP7 e *A. brasilense* FP2 e constatamos que algumas sequências mapeavam somente na estirpe Sp7, essas sequencias representavam genes que deveriam estar presentes no genoma de FP2, como por exemplo o gene *gnlZ*. Devido a esse resultado fizemos uma nova análise dos dados de genoma de FP2.

4.1.1 Remontagem do Genoma da Bactéria *Azospirillum brasilense* FP2.

A montagem da *A. brasilense* FP2 feita por GUIZELINI (2016) teve os dados de SOLiD, MiSeq e 454 com filtragem de qualidade e filtro de bases para remoção de sequências de bases ambíguas (Ns). As montagens foram realizadas no Velvet, no CLC, no Edena V3 e no MaSuRCA. A separação do cromossomo e os plasmídeos teve como referência o genoma da *A. brasilense* SP245. E foi utilizado o G-Finisher para o fechamento e a circularização do cromossomo e dos plasmídeos 1, 2 e 6. A anotação automática foi realizada no SILA e no RAST.

Para realizar a remontagem do genoma da bactéria *A. brasilense* FP2 as bibliotecas de sequenciamentos usadas foram:

Biblioteca 1: 454TrimmedReads.fasta: 128.957 *reads*;

Biblioteca 2: FP2_S5_L001_R1_001.fastq e FP2_S5_L001_R2_001.fastq: 777.652 *reads* (Illumina);

Biblioteca 3: IonXpress_001_FP2_run147.fastq: **7.788.912** *reads*;

Biblioteca 4: 413 *contigs* NCBI; Accesso: NZ_APHV01000413;

Biblioteca 5: Fasteris: *A. brasilense* FP2 – pareado (~40pb);

Bibliotecas 6: mutantes: **a)** HM014, **b)** HM053, **c)** HM210, **d)** HM26 (pareados); **e)** HM053-IonXpress;

Biblioteca 7: sequenciamento do Azospirillum brasilense Sp7.

Para realizar a remontagem do genoma da *A. brasilense* FP2 foi utilizada as bibliotecas de "dados brutos" (1, 2 e 3). O *software* CLC Genomics Workbench 7 foi usado para realizar os mapeamentos e para extrair as sequencias consensos. As versões de montagens que já haviam disponíveis foram usadas apenas para comparação entre montagens.

Inicialmente o mapeamento foi feito utilizando as bibliotecas 1 e 2 e tendo como referência o genoma da estirpe *A. brasilense* Sp7. Os *reads* não mapeados no cromossomo foram mapeados no plasmídeo 1, e assim sucessivamente até o plasmídeo 5. Os *reads* que não mapearam nos plasmídeos de 1 a 5 foram utilizados em uma montagem *de novo* (montagem sem referência), para montar o plasmídeo 6, havendo assim a estimativa de sete replicons. Os parâmetros de mapeamentos para esta montagem foram: comprimento da fração 0,5 e similaridade da fração 0,8 e os m*atch* não específicos foram mapeados randomicamente.

As sequências consensos dos sete replicons foram dadas como arquivo de entrada (*input*) no *software* FGap (Fecha Gap) juntamente com os 413 *contigs* que foram depositados no site do NCBI em 2013 (Accesso: NZ_APHV01000413). Nesses contigs há a anotação de 6.769 genes, sendo 5.755 codificadores.

A montagem resultante do FGap foi denominada Montagem 1.

Os 413 *contigs* tratam-se de uma montagem prévia do genoma do *A. brasilense* FP2, que utilizou bibliotecas diferentes das que estão sendo utilizadas aqui nesta remontagem. Sendo assim, podem possuir informações que complementam a nova montagem.

O *software* FGap teve o seguinte resultado para o cromossomo:

Antes do FGAP:

- Gaps: 32
- Sequências: 1
- Tamanho: 3.005.735bp
- GC: 68,14%
- Ns: 3.452

Depois do FGAP:

- Gaps: 7

- Sequência(s): 1
- Tamanho: 3.005.740bp
- GC: 68,18%
- Ns: 1.427

O perfil da Montagem 1 com a estimativa dos seis plasmídeos é a seguinte:

- **Plasmídeo 1:** 1.754.530pb

  2 gaps
- **Plasmídeo 2:** 819.021pb

  5 gaps
- **Plasmídeo** 3: 645.043pb

  3 gaps
- **Plasmídeo 4:** 206.692pb

  3 gaps
- **Plasmídeo 5:** 156.474pb

  2 gaps
- **Plasmídeo 6:** 815.767pb

  463 contigs

Após essa montagem foi realizado um novo sequenciamento do genoma do *A. brasilense* FP2, com a plataforma Ion Torrent, cujos *reads* possuem o tamanho de 114pb em média. Então, foi realizada uma nova montagem utilizando os dados gerados pelo Ion Torrent, denominada Montagem 2. Os parâmetros de mapeamento para a Montagem 2 foram: comprimento da fração 0,8 e similaridade da fração 0,8 e os m*atch* não específicos foram mapeados randomicamente. A tabela 01 mostra o resultado do mapeamento feito utilizando a biblioteca 3 e tendo como referência o genoma do *Azospirillum brasilense* SP7, com os plasmídeos mapeados sucessivamente, como na Montagem1. O fluxo de trabalho da Montagem 1 e da Montagem 2 está representado na Figura 03.

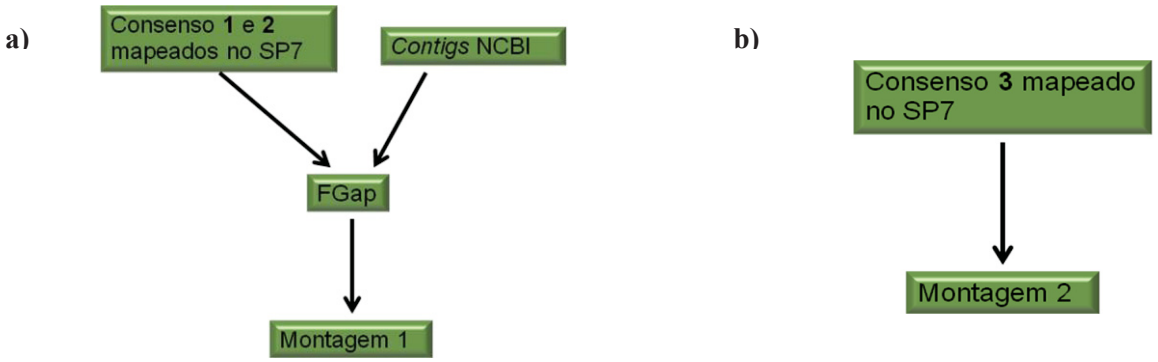


FIGURA 03: A) FLUXO DE TRABALHO PARA A MONTAGEM 1; **B)** FLUXO DE TRABALHO PARA A MONTAGEM 2.

TABELA 01: DADOS DE MAPEAMENTO DA MONTAGEM 2.

Azospirillum brasilense Sp7	***Reads* mapeados**	***Reads* não mapeados**	**Nº de gaps**
Cromossomo – 3.005.726pb	3.165.714 (40,64%)	4.623.198 (59,35%)	53 *gaps*
Plasmídeo 1 – 1.754.523pb	1.662.800 (21,34%)	2.960.398 (38,00%)	81 *gaps*
Plasmídeo 2 – 819.020pb	750.995 (9,64%)	2.209.403 (28,36%)	43 *gaps*
Plasmídeo 3 – 645.064pb	554.077 (7,11%)	1.655.326 (21,25%)	77 *gaps*
Plasmídeo 4 – 206.714pb	123.828 (1,58%)	1.531.498 (19,66%)	27 *gaps*
Plasmídeo 5 – 156.480	166.096 (2,13%)	1.365.402 (17,53%)	13 *gaps*
Plasmídeo 6	3.201 *contigs*, com tamanho médio de 323pb e 1.034.299pb total.		

Foram selecionado os 10 maiores *contigs* do plasmídeo 6 da montagem 2, com os seguintes tamanhos:

Contig_01: 55.082pb

Contig_03: 47.989pb

Contig_05: 25.988pb

Contig_06: 18.256pb

Contig_07: 27.610pb

Contig_10: 19.601pb

Contig_16: 17.626pb

Contig_18: 31.290pb

Contig_31: 19.371pb

Contig_51: 18.235pb

Para saber se os *contigs* da montagem do plasmídeo 6 realmente correspondem a um novo plasmídeo ou se pertencem à algum dos replicons anteriores foi feita uma busca no banco de dados do NCBI, através da ferramenta BLAST, foi possível ver que o contig_01 trata-se de um bacteriófago de *A. brasilense* Cd, com **99%** de identidade. O contig_02 possui 7.428pb e também tem **99%** de um bacteriófago de *A. brasilense* Cd. O bacteriofago Cd tem o tamanho de 62.337pb. A soma do contig_01 com o contig_02 é de 62.510pb, dando indício de que compõem o mesmo bacteriófago.

Ao comparar os *contigs* 01 e 02 com os *contigs* do *A. brasilense* SP7 disponibilizados no NCBI pelo grupo de Jerusalem (Hebrew University of Jerusalem), eles estão presentes no Contig_seq_16. Nos 413 contigs do *A. brasilense* FP2 disponibilizados no NCBI pelo grupo GENOPAR, eles estão presentes no id_303_Contig. Na montagem do GUIZELINI (2016), estão presentes no Plasmídeo_05.

Os demais *contigs* (Top 10) possuem similaridade com o plasmídeo 1 ou 3 do *A. brasilense* SP245.

A anotação desses 10 *contigs* possui 306 genes, e foi realizada pelo *software* SILA.

Esse resultado e demais observações dos resultados do BLAST com os outros *contigs* apontam o caminho de que não há um sexto plasmídeo, resta saber onde tais *contigs* se encaixam nos replicons. Para saber disso, será realizada uma busca das sequências desses *contigs* em outras versões genomas montados e disponíveis.

Para comparar a montagem 1 com a montagem 2 (apenas o cromossomo) foi realizado um gráfico *dotpot* com o *software* MUMmer 3.0, conforme a Figura 04. O gráfico mostra que há alta similaridade entre as montagens, havendo apenas três pequenos pontos de divergência ou de *gaps*.

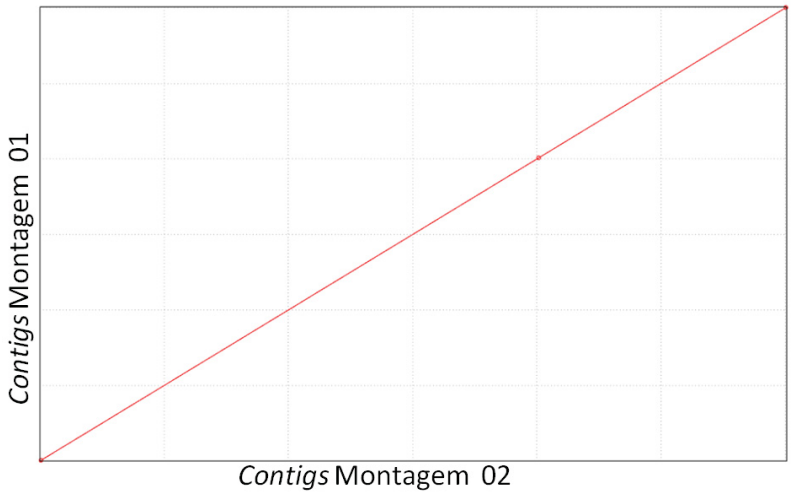


FIGURA 04: GRÁFICO *DOTPOT* COMPARANDO A MONTAGEM 1 COM A MONTAGEM 2, REALIZADO COM O *SOFTWARE* MUMMER 3.0.

Também foi realizado uma montagem *de novo* (montagem sem referência) com o *software* CLC Genomics Workbench 7 utilizando os dados da biblioteca 3, tendo como resultado 4.015 *contigs*. A Figura 05 mostra a comparação dos *contigs* da montagem de novo com o cromossomo da *A. brasilense* SP7, em que é possível ver que os *contigs* dão cobertura ao cromossomo.

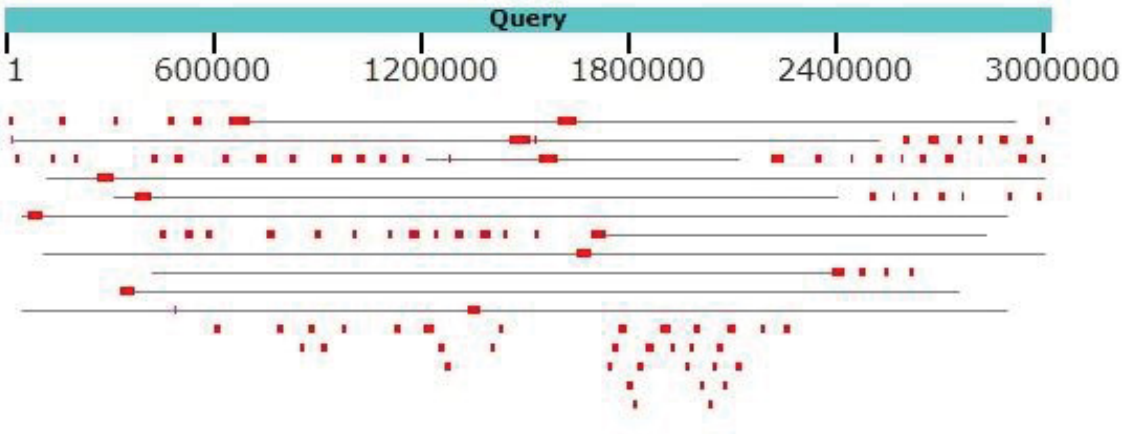


FIGURA 05: COMPARAÇÃO DA MONTAGEM *DE NOVO* COM O CROMOSSOMO DO *A. brasilense* SP7. A barra azul denominada *query* representa o cromossomo da *A. brasilense* SP7, e as barras vermelhas representam o local onde os *contigs* alinharam no cromossomo, com identidade >= 200. Gráfico produzido pela ferramenta Blast.

As demais bibliotecas (5, 6 e 7) foram utilizadas apenas para verificar se davam cobertura às regiões dos *gaps* fechados pelo FGap na montagem 1, assim validando-os.

Também foi utilizado o programa G-Finisher (GUIZELINI, 2016) com o obejtivo de melhorar a montagem. O programa inseriu mais gaps no genoma, devido ao método que o algoritmo utiliza, de quebrar *contigs* em regiões de possíveis erros de montagem e tentar reordenar os *contigs*. Portanto, o resultado do programa G-Finisher não pôde ser aproveitado.

A montagem realizada por Guizelini (2016) utilizou o genoma da *A. brasilense* SP245 como referência e depois o genoma da *A. brasilense* SP7 como referência, faltando identificar o arranjo que melhor acomode os genes. Guizelini (2016) não descartou a eventual necessidade de mais dados de sequenciamento para confirmar o número de plasmídeos, os respectivos tamanhos e organização.

Ainda se faz necessário refinar estas montagens para obter uma montagem de genoma mais completa e confiável. Uma estratégia para isso é mapear *reads* nas adjacências dos *contigs*, estendendo-os, com o objetivo de fechar *gaps*. Após a obtenção de uma nova montagem, uma

nova anotação do genoma foi realizada, utilizando o anotador SILA, foram anotados o cromossomo e os cinco plasmídeos da Montagem 01 e o cromossomo e os cinco plasmídeos da Montagem 02. Com os *reads* não mapeados nos replicons da *A. brasilense* SP7 foi realizado uma montagem *denovo* (montagem sem referência) e a anotação dessa montagem, para verificar os genes que não estão presentes em *A. brasilense* SP7.

O programa SILA anotou um total de 7.981 genes, juntando os replicons da Montagem 01, os replicons da Montagem 02 e os genes constituídos dos *reads* não mapeados nas montagens. Essa montagem foi utilizada para mapear os dados de RNAseq.

4.1.2. Teste de trimagem e de mapeamento para os transcriptomas de *Azospirillum brasilense* FP2

Para a análise dos dados de RNA-Seq de *A.brasilense* FP2 cultivada em condições de Fixação de Nitrogênio foram testadas diferentes trimagens e mapeadores.

Trimagens de cinco, dez e quinze bases no sentido 3' foram realizadas, com o objetivo de verificar qual desses trimagens otimizaria o número de *reads* mapeados. Desses três testes de trimagens, o que obteve o melhor número de *reads* mapeados foi o de 15 bases (Tabela 02). Os valores de qualidade das bases tendem a diminuir no sentido 3', e a trimagem otimiza o mapeamento quando se usa o parâmetro de qualidade de Phred 20 na média de qualidade dos *reads*.

TABELA 02: TESTES DE MAPEAMENTOS REALIZADOS COM O TRANSCRIPTOMA DE *Azospirillum brasilense* FP2.

Montagem Guizelini (2016)								Remontagem do Genoma
Biblioteca	**Dados brutos**	**CLC6 – Dados brutos**	**CLC6 – trim 5**	**CLC6 – trim 10**	**CLC6 – trim 15**	**CLC7 – trim 15**	**SHiMP trim 15**	**CLC 7 - trim 15**
Controle1	7.316.774	792.468	812.749	879.009	892.125	673.162	847.855	**1.429.697**
Controle2	13.351.128	868.789	898.396	986.460	1.011.590	764.959	982.521	**2.353.759**
N1	13.601.907	315.567	329.537	373.508	377.214	278.105	344.765	**1.031.033**
AzoN1	12.896.900	443.713	466.862	519.433	512.344	391.901	468.605	**1.291.490**

Também testamos diferentes mapeadores, CLC Workbench 6, CLC Workbench 7 e SHRiMP, para testar qual deles seria mais eficiente no mapeamento. O CLC 6 mapeou o cromossomo e os seis plasmídeos simultaneamente. CLC 7 mapeou o cromossomo e os seis plasmídeos individualmente (*tracks* de genes) e os *reads* não mapeados no cromossomo foram usados para mapear no plasmídeo 1 e assim sucessivamente. A contagem dos *reads* total foi feita no programa Microsoft Excel e o RPKM foi recalculado usando o total de *reads*, somando o Cromossomo e seis plasmídeos. No SHRiMP o mapeamento foi feito em arquivo fasta único, com cromossomo e plasmídeos. Desses, o *software* CLC Workbench 6 teve o melhor desempenho no mapeamento. Os parâmetros de mapeamento foram os mesmos para os três mapeamentos.

Segundo os resultados mostrados na tabela 2, o CLC6 com quinze de trimagem obteve os melhores resultados, por isso nossos resultados de RNAseq serão mapeados utilizando o CLC6 com quinze de trimagem.

4.1.3. Análise de qualidade das Sequências de RNA

A qualidade do sequenciamento de RNA é fundamental para a confiabilidade dos dados. As tecnologias de sequenciamento disponibilizam, juntamente com a sequência de bases, a qualidade de cada base. À medida que se estende o tamanho do *read* a qualidade decai. Ao remover as últimas bases no sentido 5', ocorre um aumento na média de qualidade dos *reads* e otimiza o número de *reads* mapeados.

O *software* FastQC faz gráficos de análise de qualidade dos *reads*. Para cada posição dos nucleotídeos um gráfico tipo 'BoxWhisker' é desenhado.

Visto que a qualidade das bases, na maioria dos sequenciadores, irá diminuir enquanto a corrida progride, é possível ver que a qualidade das bases diminui para a área em laranja e vermelha no final de uma leitura.

A Figura 06 exemplifica uma análise de qualidade dos dados de uma corrida de RNA-Seq da condição de Fixação de Nitrogênio. Os gráficos foram feitos pelo programa FastQC. Pode-se observar que a qualidade das bases diminui a medida que se aproxima do fim dos *reads* (figura 06a). E, após a trimagem das bases, a mediana do valor de qualidade de cada posição aumenta, bem como a média de qualidade de cada *read*, conforme ilustra a Figura 06b. Assim é possível otimizar o numero de *reads* mapeados, conforme mostrado na Tabela 01.

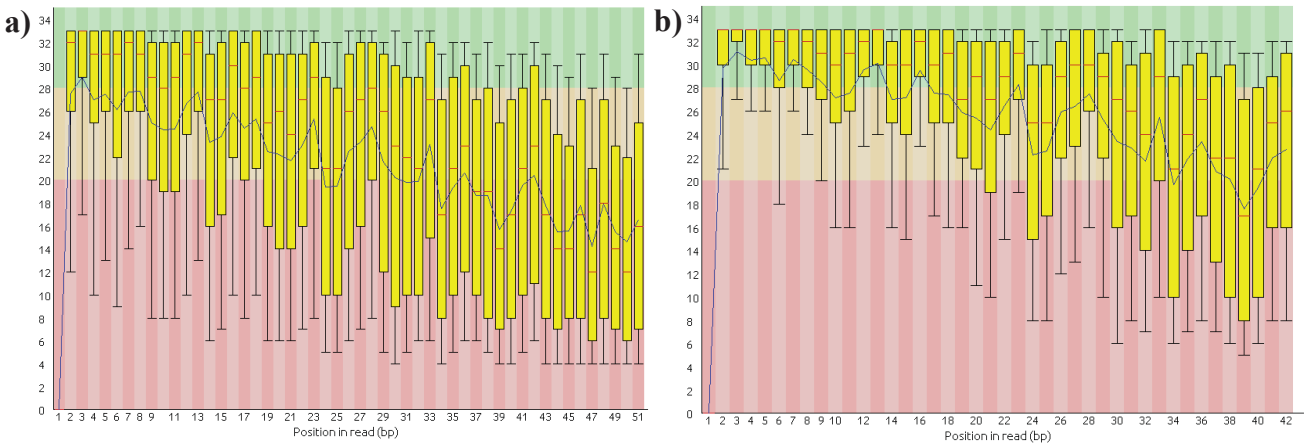


FIGURA 06: GRÁFICOS DE ANÁLISE DE QUALIDADE DA BIBLIOTECA AzoN1 (CONDIÇÃO FIXANDO - BIBLIOTECA 2).

Gráficos Box Whisker feitos com o programa FastQC. A linha vermelha central é o valor da mediana, a caixa amarela representa o intervalo inter-quartil (25-75%), os traços superiores e inferiores representam os pontos de 10% e 90%, a linha azul representa a qualidade média. O eixo y no gráfico mostra os índices de qualidade. Quanto maior o valor, melhor a qualidade da base. As cores ao fundo do gráfico divide o eixo y em camadas de muito boa qualidade (verde), camadas de qualidade razoável (laranja) e camadas de má qualidade (vermelho). a) dos dados brutos e b) dos *reads* após a trimagem de 15 bases no sentido 3'.

4.1.4. Análise de Correlação entre as amostras de RNA-Seq de *Azospirillum brasilense* FP2

A Análise de Componente Principal (PCA) é importante porque faz uma correlação entre as bibliotecas de RNA-Seq, extraindo características em comum entre elas.

Ao realizar a PCA com as cinco amostras disponíveis, a amostra AZOn2 ficou muito distante das outras, pois possui um menor número de *reads* (Figura 7). Ao retirar esta amostra da análise PCA as amostras da mesma condição biológica se correlacionaram. Portanto a biblioteca AZOn2 foi excluída das análises.

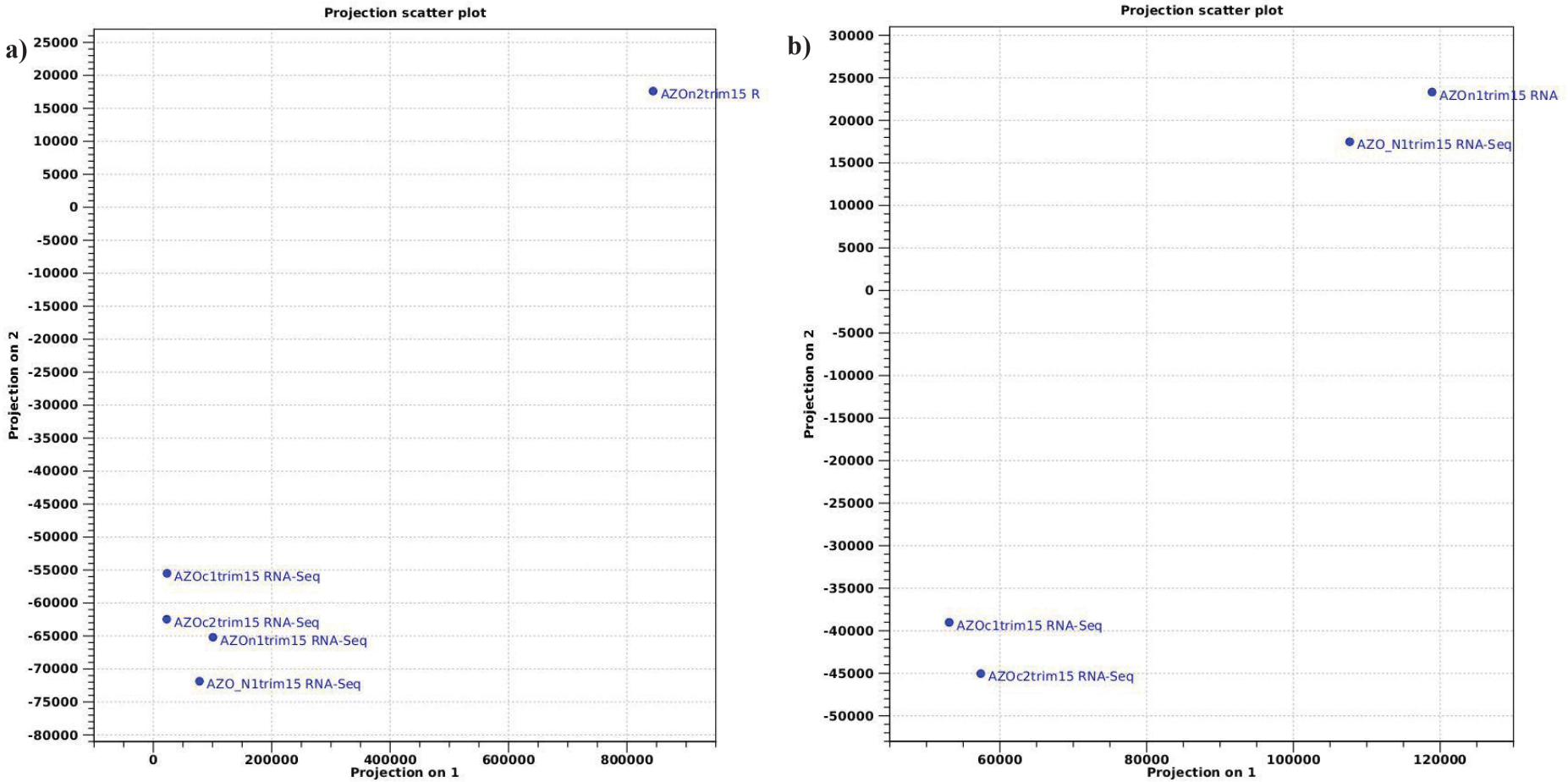


FIGURA 07: ANÁLISE DE COMPONENTE PRINCIPAL (PCA) DAS AMOSTRAS DE RNA-SEQ DE *Azospirillum brasilense* FP2.

a) Com as cinco amostras disponível: Controle1 (AZOc1trim15), Controle2 (AZOc2trim15), N1 (AZOn1trim15), N2 (AZOn2trim15) e AzoN1 (AZO_N1trim15). **b)** Com quatro amostras: Controle1 (AZOc1trim15), Controle2 (AZOc2trim15), AzoN1 (AZO_N1trim15) e N1 (AZOn1trim15).

4.1.5. Identificação dos genes diferencialmente expressos por *A. brasilense* crescido em condições de Fixação de Nitrogênio.

Para identificar os genes Diferencialmente Expressos (DE) por *A. brasilense*, crescidos em condições de fixação de nitrogênio, foram utilizados três testes estatísticos: o Baggerley's test no *software* CLC Worchbench, o DESeq – BioConductor – R e o EdgeR – BioConductor – R. Foram selecionados os genes com *fold-change* maior ou igual a 2 para os ativados; e os genes com *fold-change* menor ou igual a -2 para os reprimidos. E com o *p-value* menor que 0,05. O resultado dos três métodos estatísticos estão na Tabela 03.

TABELA 03: GENES DIFERENCIALMENTE EXPRESSOS DE ACORDO COM TRÊS PROGRAMAS DE ANÁLISE ESTATÍSTICA.

Métodos	Número de genes D.E.	UpRegulados	DownRegulados
CLC-Baggerley's test	2.623	243	2.380
EdgeR – BioConductor – R	1.905	773	1.132
DESeq – BioConductor – R	1.569	686	883
Total	**3.234**	**822**	**2.412**
Comum aos Métodos	**1109**	**229**	**880**

A Tabela 03 mostra a quantidade de genes diferencialmente expressos para cada um dos métodos estatísticos realizados e quantos desses genes estão ativados ou reprimidos em cada método. A linha Total mostra o total de genes diferencialmente expressos únicos, somando o resultado dos três métodos. A diferença da quantidade de genes diferencialmente expressos em cada programa ocorre porque implementam métodos estatísticos ou modelos matemáticos distintos. O Baggerley's test contido no CLC pondera os valores de expressão gênica entre as replicatas, tendo como resultado menos rigoroso em relação os outros métodos utilizados. Pode-se observar que o Baggerley's test obteve o maior número de genes diferencialemte expressos (1.714 genes) entre os três programas. O EdgeR (BioConductor – R) focaliza na expressão diferencial em relação ao tamanho das bibliotecas, o que tende a minimizar o *fold-change*. De acordo com os dados mostrados na Tabela 03, obteve menos genes D.E. que o Baggerley's test. Já o DESeq (BioConductor – R) gera um fator de normalização com o tamanho relativo das

bibliotecas e o valor em torno da média, multiplicando cada valor de expressão gênica por esse fator, obtendo assim um resultado mais rigoroso. Sendo assim, os gráficos mostrados neste trabalho utilizam os valores de log2fold-change obtidos pelo DESEq. Mas, optou-se por analisar apenas o genes D.E. de acordo com a intersecção dos resultados.

Os diagramas de Venn na Figura 08 mostra a intersecção dos genes diferencialmente expressos nos três testes estatísticos aplicados (a) e quantos desses genes estão ativados nos três testes simultaneamente (b) ou reprimidos (c).

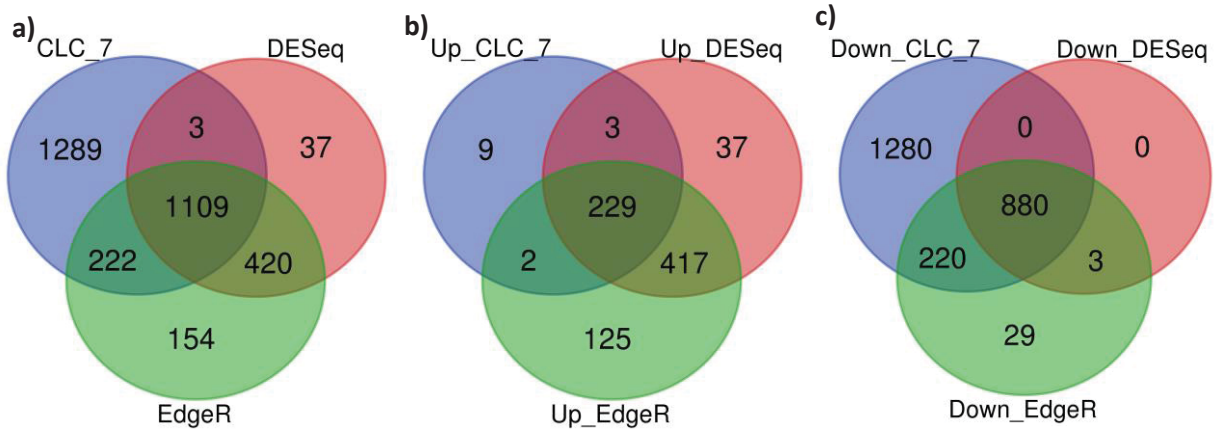


FIGURA 08: DIAGRAMA DE VENN COM A QUANTIDADE DE GENES DIFERENCIALMENTE EXPRESSOS DE ACORDO COM OS MÉTODOS BAGGERLEY'S TEST NO *SOFTWARE* CLC WORCHBENCH, O DESEQ – BIOCONDUCTOR – R E O EDGER – BIOCONDUCTOR – R.

a) Diagrama de Venn de todos os genes diferencialmente expressos (DE). **b)** Diagrama de Venn dos genes DE ativados **c)** Diagrama de Venn dos genes DE reprimidos. Diagramas feitos pelo site http://bioinformatics.psb.ugent.be/webtools/Venn/.

Foram considerados diferencialmente expressos os genes cuja diferença de expressão (fold-change) foi maior que dois ou menor que -2 e cujo valor de p foi menor que 0,05. O Anexo 01 contem a lista dos 880 genes reprimidos e o Anexo 02 contem a lista dos 229 genes ativados, com os valores de log2 fold-change e p-value dos três programas de estatística utilizados.

Foi obtido o código da função do gene com a ferramenta KAAS (KEGG Automatic Annotation Server), no qual é colocado um arquivo com as sequências de nucleotídeos dos genes DE e é retornado o código KO (KEGG Orthology) dos genes que encontra alguma similaridade no banco de dados (MORIYA, 2007). O conjunto de genes diferencialmente expressos foi utilizado para alimentar a ferramenta KAAS, buscou-se saber em que vias metabólicas tais genes estão envolvidos utilizando a ferramenta KEGG (Kyoto Encyclopedia of Genes and Genomes).

Tendo em mãos o código KO dos genes DE pode-se então verificar em quais vias metabólicas estão participando.

A Figura 09 mostra a expressão diferencial desses genes, na escala log2 a partir dos dados do programa DESeq, focando nas principais vias metabólicas.

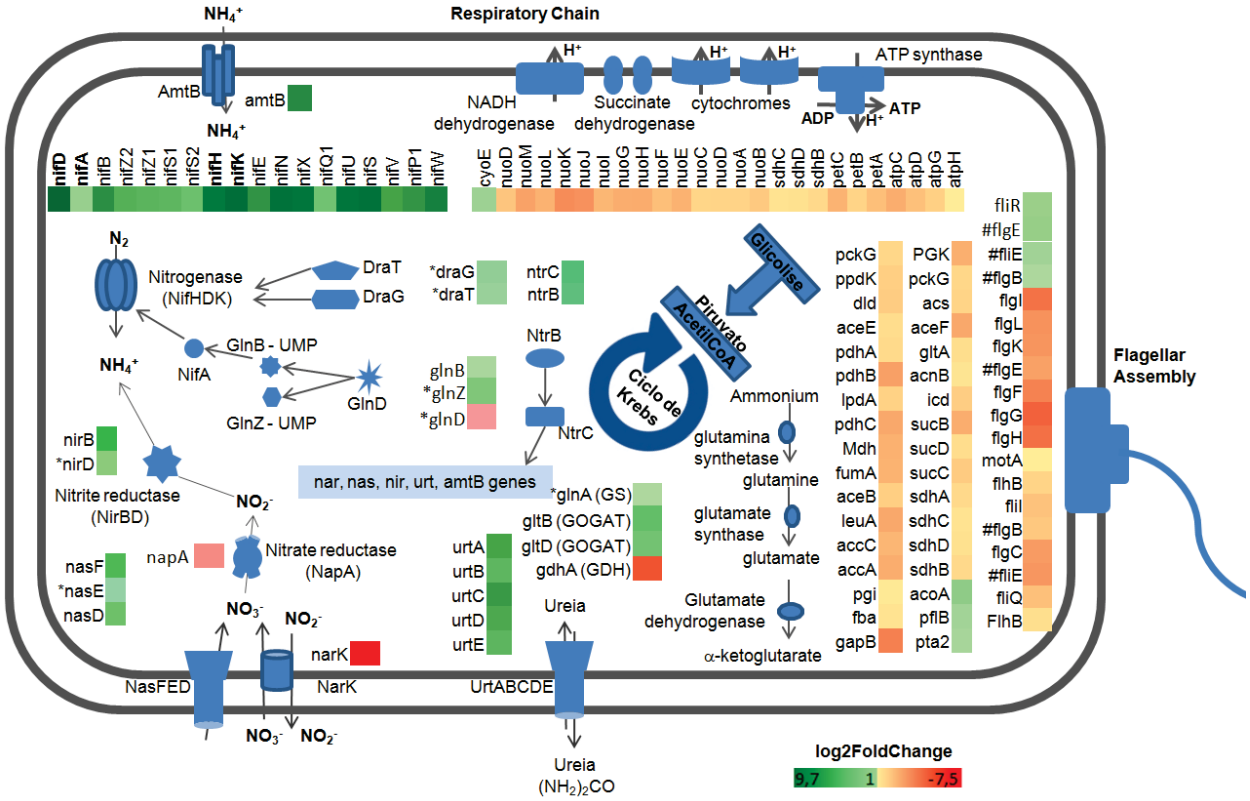


FIGURA 09: PANORAMA DAS VIAS METABÓLICAS NA BACTÉRIA *Azospirillum brasilense* FP2 DE ACORDO COM OS GENES DIFERENCIALMENTE EXPRESSOS EM ESCALA LOG2-FOLD-CHANGE.

Os valores de log2-fold-change foram fornecidos pelo programa DESeq. Os genes ativados estão em tons de verde ■ e os genes reprimidos estão em tons de vermelho ■. Os genes marcados com sustenido (#) possuem cópias no genoma, um *up* e o outro *down*. Os genes marcados com asterisco (*) não aparecem na intersecção dos três *softwares* de análise estatística.

4.1.6 Genes envolvidos no Metabolismo de Nitrogênio em *A. brasilense* FP2

Entre os genes ativados estão os genes *nif*, genes que codificam para enzimas envolvidas na Fixação Biológica de Nitrogênio (FBN) e cujos valores de *fold-change* estão mostrados na Figura 10.

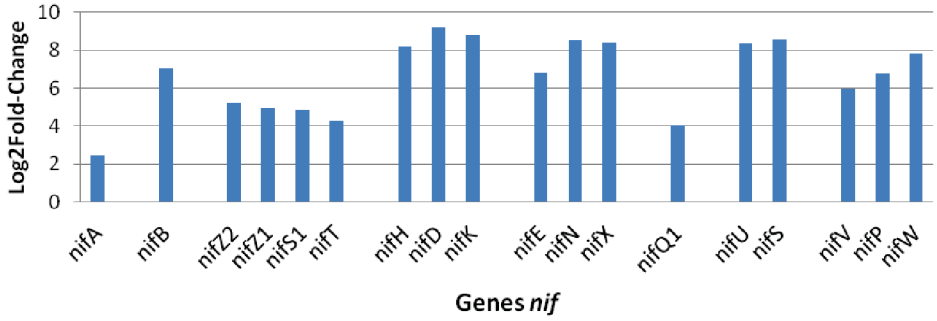


FIGURA 10: VALORES DE LOG2-FOLD-CHANGE DOS GENES *nif* ORGANIZADOS EM OPERONS.

Os valores utilizados são do programa DESEq e possuem *p-value* menor que 0,05.

Em *A. brasilense* FP2 o gene *nif*H está menos expresso que os genes *nif*DK. Esse resultado também foi verificado durante o estudo do transcriptoma de *Azotobacter vinelandii* e em *Paenibacilus sp*. WLY78 crescidos em condição de Fixação de Nitrogênio (HAMILTON *et al.,* 2011; WANG *et al*., 2016). Em *Azoarcus sp.* strain BH72 os genes *nifH* e *nifK* não aparecem diferencialmente expressos, de acordo com os parâmetros estabelecidos (SARKAR & REINHOLD-HUREK, 2014). Outros genes *nif*, que também estão envolvidos no processo de Fixação de Nitrogênio aparecem diferencialmente expressos. A Figura 11 mostra os níveis de expressão de genes *nif* que também estão diferencialmente expressos em *Azoarcus sp.*, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78. É possível observar uma grande diferença entre os níveis de expressão nos 4 organismos, que pode se justificar pelo tamanho da biblioteca de cada um e pelas peculiaridades de cada organismo.

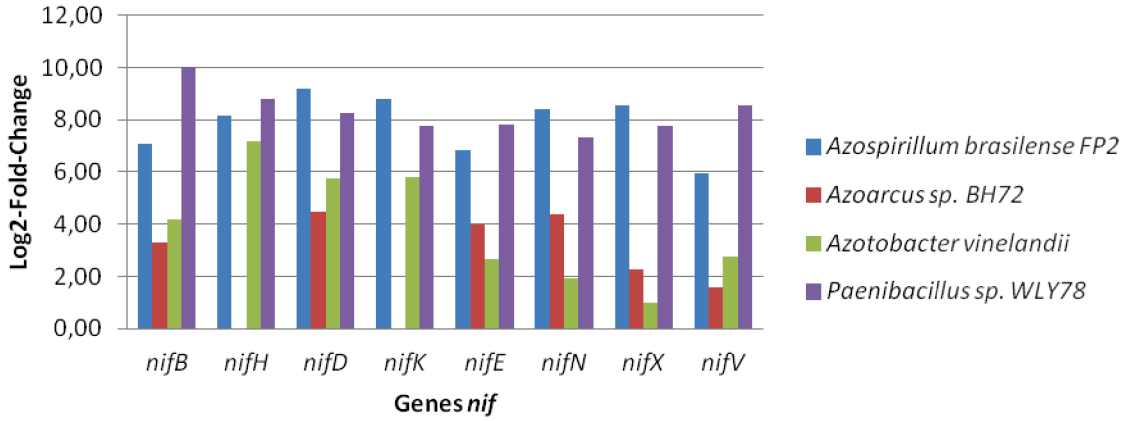


FIGURA 11: COMPARAÇÃO DOS VALORES DE LOG2-FOLD-CHANGE DOS GENES *nif* DAS BACTÉRIAS *A. brasilense* FP2, *Azoarcus sp.* STRAIN BH72, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78.

O Anexo 03 contem a comparação dos genes diferencialmete expressos da *Azospirillum brasilense* FP2 com os genes diferencialmente expressos de *Azoarcus sp*. BH72, *Azotobacter vinelandii e Paenibacillus sp.* WLY78 em condições de Fixação Biológica de Nitrogênio.

Outros genes que estão envolvidos no metabolismo de Nitrogênio aparecem ativados, como os transportadores de Amônio (*amt*B), de Amônia e Nitrato (*nas*FED e NarK) e de Uréia (*urt*ABCDE); o gene que codifica para nitrato redutase (*nap*A) aparece reprimido, enquanto os genes que codificam para a nitrito redutase (*nir*BD) aparecem ativados, assim como os genes *ntr*B, *ntr*C, e *gln*B.

A proteína uridil-transferase (GlnD) uridilila as proteínas GlnB (proteína reguladora de nitrogênio P-II) e GlnZ (DE ZAMAROCZY *et al.,*1996; DE ZAMAROCZY *et al.,*1998;), em condições de limitação de Nitrogênio. Após ser uridililada a proteína GlnB (GlnB-UMP) controla a atividade da proteína NifA (ARAUJO *et al.,* 2004), que por sua vez ativa a transcrição dos genes *nifHDK*, que ao serem traduzidos, compõem o complexo enzimático nitrogenase (VAN DOMMELEN *et al*., 2002). Conforme ilustrado pela Figura 09, os genes *gln*B e *glnZ* aparecem ativados; e deveriam estar sendo uridililados em condições de Fixação Biológica de Nitrogênio. Porém o gene que codifica para a proteína GlnD não encontra-se diferencialmente expresso (de acordo com ao parâmetros estabelecidos), com o valor de fold-change de 0,7 e cobertura de 1,64.

A uridililação de proteínas PII (GlnB e GlnZ) regula a atividade das proteínas NtrB e NtrC. GlnB, quando não modificada, interage com NtrB, uma histidina auto-quinase, induzindo sua atividade fosfatase que desativa o ativador transcricional NtrC-P por desfosforilação (BUENO *et al*., 1985; JIANG & NINFA, 1999). Em condições de limitação de $NH4^+$, GlnB uridililada não consegue interagir com NtrB, que por seguinte, ativa a NtrC por fosforilação. NtrC fosforilada (NtrC-P) interage com o complexo enzimático RNA-polimerase e ativa a transcrição de vários genes regulados por Nitrogênio: genes *nar*, *nas, nir*, *urt* e *amtB* (ATKINSON *et al.*, 1994). Os genes *nas*FED, *nir*BD e *urt*ABCDE estão organizados em operons e ativados em *A. brasiliense* FP2 conforme ilustrado na Figura 09. O gene *nirB* também se encontra ativado em *Paenibacilus sp*. WLY78, com fold-change de 3,02 e reprimido em *Azoarcus sp.* strain BH72, com fold-change de -2,1. Os genes *nasED* também estão ativados em *Paenibacilus sp*. WLY78.

Em *Azoarcus sp.* Strain BH72 os genes *nas*FD e *urt*E estão regulados no mutante *nif*LA^-, indicando que *nif*A atua como um ativador transcricional para genes que participam da Fixação Biológica de Nitrogênio (SARKAR & REINHOLD-HUREK, 2014).

Em *A. brasilense*, os íons de amônio derivados da fixação do nitrogênio ou do ambiente, são assimilados pela ativação da glutamina sintetase (GS) e glutamato sintase (GOGAT), independente do nível de Nitrogênio na célula (GAUTHIER & ELMERICH, 1977; ELMERICH *et al.,* 1977; PELANDA *et al*., 1993). Uma outra maneira é pela ativação da glutamato dehidrogenase (GDH) (OKON *et al*., 1976; WESTBY *et al*., 1987). O perfil do transcriptoma aponta que a via utilizada foi a GS – GOGAT, visto que a enzima GDH está reprimida. Em *Paenibacillus sp.* WLY78, uma bactéria gram-positiva, os genes que codificam para GS,

GOGAT e GDH encontram-se ativados durante a Fixação Biológica de Nitrogênio (SHI *et al*., 2016).

A atividade da nitrogenase da *A. brasilense* é inativada por amônio ou oxigênio. Esta inativação envolve ADP-ribosilação da proteína dinitrogenase redutase (NifH) catalisada pela dinitrogenase-redutase ADP-ribosiltransferase (DraT) e é revertida, após o esgotamento de NH_4^+, pela glicohidrolase dinitrogenase redutase (DraG) (ZHANG, *et al.*, 1997). A figura 09 mostra que esses dois genes estão levemente ativados em condições de fixação de nitrogênio. Mas esta ativação não é estatísticamente significativa (Os valores de fold-change são de 1,0 e 1,4 para *dra*T e *dra*G respectivamente).

4.1.7 Genes que codificam para proteínas ribossomais de *A. brasilense*

As proteínas ribossômicas são proteínas que constituem os ribossomos. As proteínas ribossômicas podem ser estruturalmente importantes dentro do ribossomo; por exemplo, podem interagir com o rRNA e podem também funcionar diretamente na tradução de mRNA em proteína, que é a principal função dos ribossomos.

O genoma da *A. brasiliense* FP2 possui 60 proteínas ribossomais anotadas. Quarenta e cinco (45) genes que codificam para proteínas ribossomais estão reprimidos durante a Fixação Biológica de Nitrogênio, conforme ilustra a Figura 12.

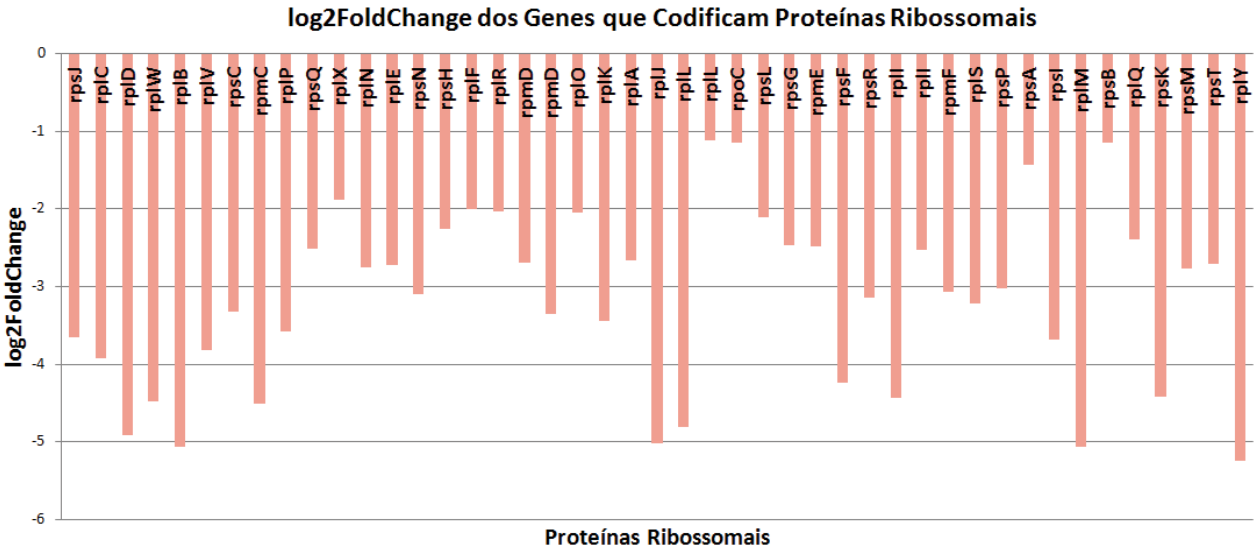


FIGURA 12: VALORES DE LOG2 FOLD-CHANGE DOS GENES QUE CODIFICAM PARA PROTEÍNAS RIBOSSOMAIS.

Em *Azoarcus sp.* BH72 alguns genes que codificam para proteínas ribossomais também aparecem reprimidos, assim como genes que codificam para o metabolismo energético (via glicolítica e ciclo do ácido cítrico), sugerindo que o alto custo energético (ATP/ADP) que ocorre durante a Fixação de Nitrogênio inibe algumas das enzimas e também sua expressão ou esse resultado pode refletir que há uma taxa de crescimento lenta (SARKAR & REINHOLD-HUREK, 2014).

Em *Azoarcus sp.* BH72 há dados de qRT-PCR sob condição similar que confirmam a repressão de genes que codificam para proteínas ribossomais (SARKAR & REINHOLD-HUREK, 2014).

4.1.8 Metabolismo energético de *A. brasilense* crescido em condições de Fixação de Nitrogênio

A figura 13 compara a diferença de expressão (*fold-change*) de genes que participam da Fosforilação Oxidativa de *A. brasilense* FP2, *Azoarcus sp.* BH72, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78. Para *Azoarcus sp.* BH72 a seleção dos genes diferencialmente expressos foram de acordo com um teste-t pareado com correção de Bonferroni com pelo menos 1,8 de fold-change e p-value menor ou igual a 0,05 (SARKAR & REINHOLD-HUREK, 2014). Para *Azotobacter vinelandii* a análise estatística foi realizada pelo DESEq e considerado diferencialmente expresso os genes com fold-change maior ou igual a 2 e p-value menor que 0,01 *nif*DK (HAMILTON *et al.,* 2011). Para *Paenibacilus sp*. WLY78 foi usado o programa DEGseq e foram considerados diferencialmete expressos os genes com p-value ajustado menor que 0,001 (SHI *et al*., 2016).

Os genes de *A. brasilense* FP2 que codificam para as proteínas envolvidas com a fosforilação oxidativa (ou Cadeia de Transporte de Elétrons), Ciclo do Ácido Cítrico (ou Ciclo de Krebs) e para a via da Glicólise estão reprimidos em condições de Fixação de Nitrogênio.

Alguns dos genes que participam dessas vias (*atp*C, *atp*B, *sdh*D, *sdh*A, PDHB e PDHA) de *Azoarcus sp.* BH72 também encontram-se reprimidos, enquanto que os genes de *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78 envolvidos nessa via encontram-se ativados.

Os genes *nuo* (*nuo*ABCDEFGHIJKLMN), que compõem o complexo da NADH dehidrogenase na cadeia de transporte de elétrons, estão reprimidos em *A. brasilense* FP2;

enquanto que em *Azotobacter vinelandii* encontram-se ativados em condição de Fixação Biológica de Nitrogênio.

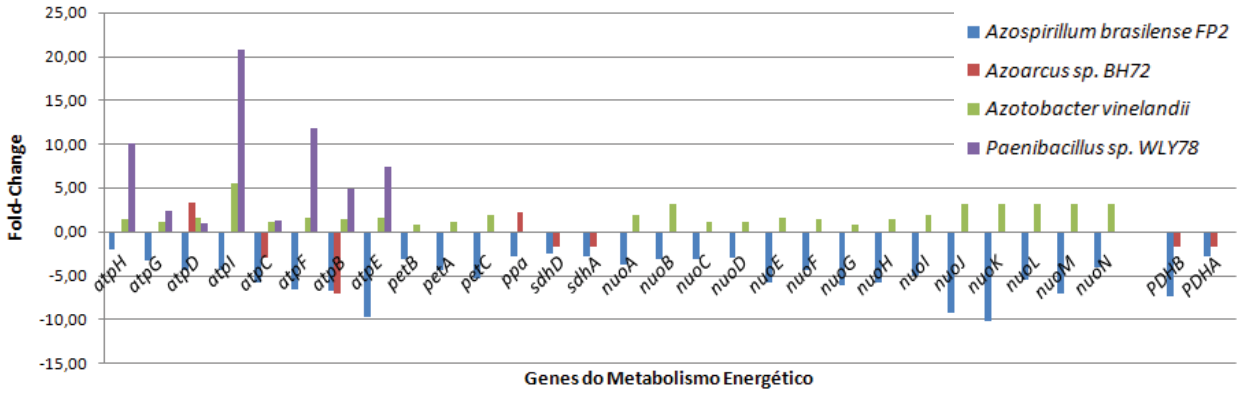


FIGURA 13: COMPARAÇÃO DOS VALORES DE *FOLD-CHANGE* DOS GENES QUE CODIFICAM PARA PROTEÍNAS QUE PARTICIPAM DA DO METABOLISMO ENERGÉTICO em *A. brasilense* FP2, *Azoarcus sp.* BH72, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78.

De acordo com Sarkar & Reinhold-Hurek (2014), genes particularmente relacionados ao metabolismo energético foram encontrados reprimidos sob N_2, talvez correlacionando com o abrandamento do metabolismo geral associado ao momento de Fixação de Nitrogênio.

Diferentemente de *A. brasilense* FP2, uma alfaproteobacteria, *Azotobacter vinelandii*, uma gammaproteobacteria, tem alguns genes que participam do metabolismo energético ativados. *A. vinelandii* possui três tipos de Nitrogenases, uma sendo expressa na presença de Molibdênio (Mo), outra na presença de Vanádio (V) e a terceira é expressa apenas na ausência de Mo e V (JOERGER *et al.,* 1989). Diferenças como estas podem justificar as diferenças de metabolismo durante a Fixação Biológica de Nitrogênio dessas bactérias.

4.1.9 Genes Flagelares de *Azospirillum brasilense*

O flagelo bacteriano é uma estrutura supramolecular complexa, que compreende um sistema de transporte de proteínas e uma máquina molecular que permite o movimento celular bacteriano (OSTERMAN *et al.*, 2015). A expressão dos genes que participam da montagem flagelar é regulada nos níveis transcricionais, traducionais e no transporte.

Cada flagelo consiste de um corpo basal, um filamento externo longo feito de proteína flagelina e um "gancho" que liga essas duas partes (CHEN *et al.,* 2011). O flagelo rotaciona e a energia para isso é fornecida pela transmembrana (BLAIR & BERG, 1990). Além da função motora, o flagelo possui a função de transporte que é usada durante o elasteamento (IINO, 1969). O transporte de proteínas compreendendo a parte extracelular do flagelo ocorre através do eixo central desta estrutura (JONES & AIZAWA, 1991).

Mais de sessenta proteínas diferentes participam da montagem e funcionamento do flagelo (OSTERMAN *et al.*, 2015). O genoma da *A. brasilense* FP2 possui mais de 71 genes flagelares anotados (contando as cópias), dentre eles estão: *flg*D, *flg*I, *fli*X, *flg*H, *flg*G, *flg*F, *fli*L, *fli*M, *mot*A, *fli*P, *fhl*B, *flg*B, *flg*C, *fli*E, *fli*Q, *fli*R, *flh*B, *fli*G, *fli*N, *fli*F, *flg*K, *fli*S, *fli*D, *flg*L, *flg*E, *flh*A. A análise do transcriptoma revelou que 24 desses genes estão reprimidos em condição de Fixação de Nitrogênio (Figura 14).

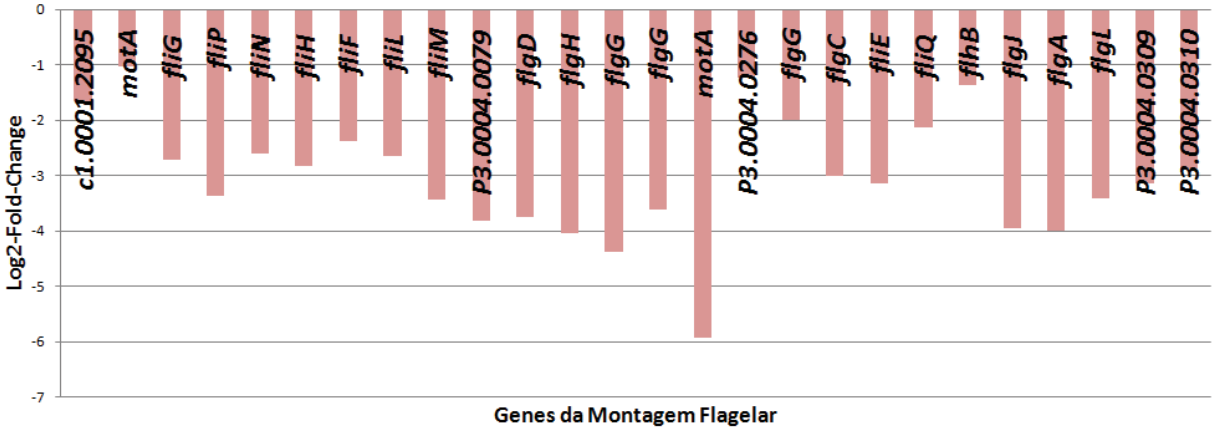


FIGURA 14: GENES DA VIA DA MONTAGEM FLAGELAR QUE ESTÃO REPRIMIDOS em *A. brasilense* FP2 de acordo com os três programas de análise estatística e dentro dos parâmentros de seleção. Os valores de *fold-change* estão em escala de log2.

A ativação da transcrição dos genes flagelares depende da expressão dos genes *flh*D e *flh*C. As proteínas FlhD e FlhC formam o fator transcricional FlhDC e somente juntos eles podem iniciar a cascata de expressão dos genes flagelares. A ativação da transcrição ocorre pela interação entre FlhDC e a subunidade α da RNA polimerase. O fator transcricional FlhDC ativa genes que são transcritos pela RNA polimerase contendo o fator σ^{70}; e que participam da montagem do inicio do flagelo, que são os componentes estruturais dos anéis, corpo basal e gancho, substratos iniciais do transporte flagelar e seus acompanhantes (KUTSUKAKE *et al.*, 1990; WANG *et al*. 2006; CHILCOTT & HUGHES, 2000; LIU *et al.*, 1995; LIU *et al.,* 1994).

Os genes *flh*C e *flh*D não estão dentro do conjunto de genes diferencialmente expressos pela intersecção dos três programas de estatítica, dentro dos parâmetros estabelecidos. No

programa DESeq o gene *flh*C está com o valor de fold-change de -1,79 e o gene *flh*D com o fold-change de -1,26.

O fator σ^{28} codificado pelo gene *fli*A e seu inibidor FlgM são importantes reguladores da montagem flagelar. A transcrição de genes tardios da montagem ocorre com a participação de σ^{28} ao invés de σ^{70} (ARNOSTI & CHAMBERLIN, 1989; OHNISHI *et al.*, 1992). Porém, o gene *fli*A não foi encontrado nos genomas das espécies de *Azospirillum*, sugerindo que outro gene, não homologo a *fli*A, pode ser o regulador. Já o inibidor *flg*M possui fold-change de -2,19.

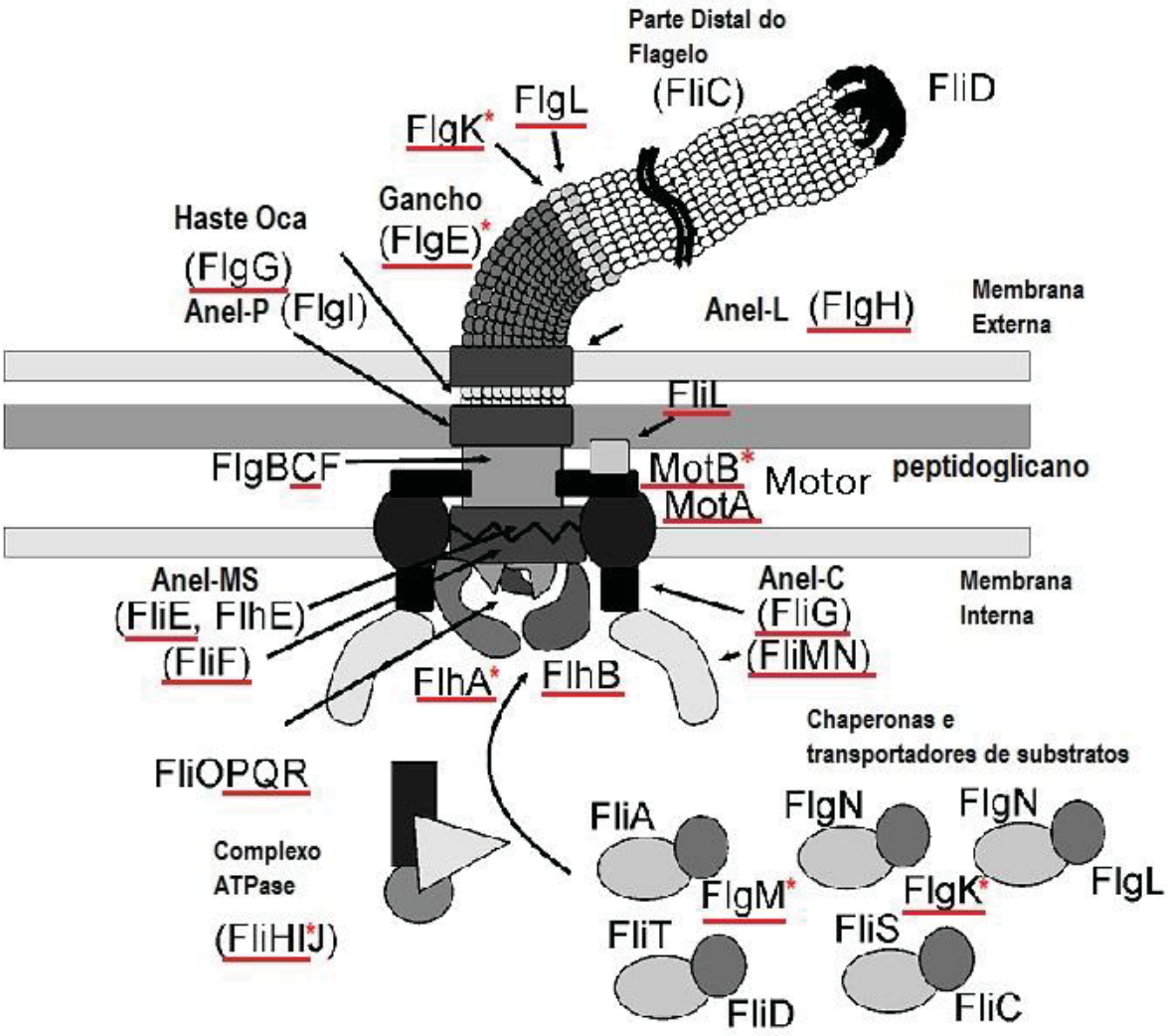


FIGURA 15: ESTRUTURA DA MONTAGEM FLAGELAR.

Ilustração das camadas de membrana interna e externa, parede celular de peptidoglicano, anéis C, MS, P e L - "rolamentos", proteínas motoras (MotAB), rotor (FlgBCF), haste oca, gancho e parte distal do flagelo. Componentes do aparato de transporte de flagelo, complexo de ATPase, chaperons de transporte flagelar e seus substratos também estão ilustrados. As proteínas sublinhadas de vermelho possuem os genes reprimidos no transcriptoma da *A. brasilense* FP2 em condição de Fixação de Nitrogênio; o asterisco em vermelho sinaliza os genes que estão reprimidos, mas não estão dentro dos parâmetros estatísticos estabelecidos. **FONTE**: adaptado de OSTERMAN *et al*., 2015.

A Figura 15 ilustra a estrutura da Montagem Flagelar dando destaque aos genes reprimidos no transcriptoma da *A. brasilense* FP2 em condição de Fixação de Nitrogênio, que estão sublinhadas de vermelho.

A expressão dos genes tardios é iniciada somente após o corpo basal e o gancho serem formados. Nos estágios iniciais da montagem do flagelo, σ^{28} é inativo, pois está localizado em um complexo com FliA/FlgM (SORENSON *et al.*, 2004), que é incapaz de interagir com a RNA polimerase. Em *A. brasilense* AZ39 os genes *flh*D e *flh*C estão presentes no plasmídeo 2. Em *A. brasilense* FP2 não foi possível identificar em qual plasmídeo se encontram, pois estão dentre os genes não mapeados no genoma da *A. brasilense* SP7.

As células de *A. brasilense* FP2 foram cultivadas em meio com NFbHPN-lactato durante 16 h a 30°C e 120 rpm e depois inoculadas por perfuração das placas de motilidade com palitos de dente estéril embebidos na cultura bacteriana. As placas utilizadas para realizar esse experimento continham meio NFbHP-lactato contendo 1% de LB, e 0,25 g/L de agar, na presença ou ausência de 20 mM de NH_4Cl^+. Os halos de crescimento foram medidos com um paquímetro após 24h de incubação a 30°C, foram feitos dois ensaios independentes, com seis repetições cada. As células de *A. brasilense* FP2 mostraram uma motilidade 4 vezes superior na presença de amônio (18,2 ± 0,2 mm de halo de crescimento) do que na ausência de amônio (4,8 ± 0,2 mm).

O conhecimento da montagem do flagelo ainda não está completamente elucidado. A regulação da tradução de componentes de mRNA do sistema flagelar ainda não foi estudada (OSTERMAN *et al*., 2015). E ainda não há descrição de como a Fixação Biológica de Nitrogênio afeta a montagem flagelar.

Esse resultado não foi encontrato em dados de transcriptoma de *Azoarcus sp.* BH72, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78 durante a Fixação Biológica de Nitrogênio, pois nestes organismos os genes da Montagem Flagelar não estão regulados.

Esse resultado indica que *A. brasilense* FP2 tem sua mobilidade reduzida durante a FBN. O FlgM é o sensor mais importante do transporte flagelar. FlgM posterga a síntese de flagelina (*fli*C) na célula até o momento em que a possibilidade de sua exportação e incorporação em um flagelo é criada (ZASLAVER *et al*., 2004).

FliA regula a transcrição de genes tardios da montagem flagelar, que são *fli*C e *fli*D. Alterações na expressão de *fli*A afeta a transcrição de *fli*C e *fli*D, justificando a redução no

tamanho do flagelo. Porém, não foi encontrado um gene homolgo ao gene *fli*A nos genomas das espécies de *Azospirillum*.

PANKIEVICZ *et al.* (2016) mostrou que o principal conjunto de genes que compõe o flagelo de *H. seropedicae* foram reprimidos ao colonizar a raiz de trigo. O gene fliP (Hsero_2049), flgK (Hsero_2042) e motB (Hsero_2985) que codificam para proteínas estruturais flagelares estavam reprimidas. Ademais, *H. seropedicae* cultivada na presença do flavonóide naringenina apresentou repressão de genes flagelares (TADRA-SFEIR *et al.* 2015).

4.1.10 Proteínas Reguladoras de transcrição de *Azospirillum brasilense*

Sete genes que codificam para proteínas reguladoras de transcrição foram encontrados ativados em condição de limitação de Amônio. A Tabela 04 contem a relação dos ativadores transcricionais e seus respectivos fold-change. Os demais reguladores transcricionais não ficaram ativados de acordo com os critérios estabelecidos.

TABELA 04: REGULADORES TRANSCRICIONAIS ATIVADOS EM *A. brasilense* FP2 EM CONDIÇÃO DE FIXAÇÃO BIOLÓGICA DE NITROGÊNIO.

ID	Fold-Change	Descrição
c1.0001.0172	4,74	Fur family transcriptional regulator
c1.0001.0979	17,51	transcriptional regulator, BadM/Rrf2 family
P1.0002.0724	6,73	Rrf2 family transcriptional regulator
c1.0001.1474	5,52	Fis family transcriptional regulator
P1.0002.1132	7,61	LysR family transcriptional regulator
P2.0003.0092	10,89	putative two component transcriptional regulator
P3.0004.0179	28,41	ArsR family transcriptional regulator

O regulador transcricional da família Fur regula genes que participam da absorção de Ferro pela célula. É um repressor sensível ao ferro que controla a expressão de genes para a biossíntese de sideróforos e o transporte de ferro. Embora Fur seja comumente considerado como um repressor dependente de metal, Fur também ativa a expressão de muitos genes por

mecanismos indiretos ou diretos. Nos sistemas de modelos melhor estudados, Fur funciona como um regulador global da homeostase do ferro controlando tanto a indução de funções de absorção de ferro (sob limitação de ferro) quanto a expressão de proteínas armazenadoras de ferro e enzimas que utilizam ferro, sob suficiência de ferro (LEE & HELMANN, 2007). Os transportadores de Fe em *A. brasilense* (*afu*ABC) não se encontram diferencialmente expresos durante a FBN.

A família Rrf2 de fatores de transcrição compreende a família de reguladores de bactérias que controlam diversos processos metabólicos, incluindo biogênese de cluster de FeS (IscR), sistemas de desintoxicação de óxido nítrico (NsrR), ferro absorção (RirA) [17173478], metabolismo de cisteína (CymR). A atividade de ligação ao DNA dos reguladores da família Rrf2 é controlada por diversos mecanismos (GIEL *et al*., 2006; RODIONOV *et al*., 2005; RODIONOV *et al*., 2006; MIDORIKAWA *et al*., 2009).

Reguladores transcricionais da família Fis ativa a transcrição de RNA ribossômico e outros genes. Desempenha um papel direto na ativação a montante dos promotores de rRNA. Também impede o início da replicação do DNA a partir de oriC e se liga a centenas de sítios de transcrição ricos em AT transcricionalmente ativos e inativos. Aproximadamente metade de seus sítios de ligação estão em DNA não-codificador (ROSS *et al*., 1990; WOLD *et al*., 1996; GRAINGER *et al*., 2006).

LysR ativa a transcrição do gene lysA, que codifica uma diaminopimelato descarboxilase. LysR é também um regulador negativo da sua própria expressão. Diaminopimelato descarboxilase é uma enzima que catalisa a clivagem de ligações carbono-carbono para produzir CO_2 e L-lisina, um aminoácido essencial (STRAGIER & PATTE, 1983).

Também encontra-se ativo um regulador transcricional do Sistema de 2 Componentes, que comprende uma série de genes, dentre eles, genes que participam da Fixação biológica de Nitrogênio, como a *nif*A.

ArsR é um repressor de transcrição para o operon *ars*EFG e é uma proteína reguladora trans-atuante que controla sua própria expressão. Os genes *ars*EFG ajudam a processar moléculas que contêm sulfato (DIORIO *et al.,* 1995).

4.1.11 Comparação dos dados de transcriptoma e proteoma de *A. brasilense*

Os dados de RNA-Seq de *A. brasilense* FP2 crescido em condições de Fixação de Nitrogênio e em alto Amônio foram comparados com os resultados de proteômica da *A. brasilense* FP2 crescidos em semelhantes condições (KADOWAKI *et al.*, 2011; PEDROSA *et al.*; 2010).

Os resultados de proteômica foram realizados por Marina Kadowaki e confirmaram alguns dados do transcriptoma. As informações do proteoma são análises dos géis 2D realizadas pelo *software* ImageMaster, em que são gerados a porcentagem de volume de cada *spot* de proteína. A porcentagem de volume está relacionado a intensidade de coloração dos *spots* por azul de coomassie. Somente a diferença de intensidade de porcentagem de volume de coloração dos *spots* não é suficiente para inferir diferença de expressão em proteômica, porque tem proteína que expressa pouco (baixa intensidade) mas é imprescindível para o metabolismo e ela não exibirá essa diferença de intensidade. As proteínas foram identificadas pelo programa MASCOT MS/MS.

A análise dos *spots* presentes nos géis contendo o extrato de bactérias crescidas em condição de Fixação de Nitrogênio mostrou que algumas proteínas (6) eram expressas somente em condições de fixação de nitrogênio. A expressão dos genes que codificam para essas proteínas está aumentada no transcriptoma (Tabela 05).

O anexo 04 contém a lista de todas as proteínas identificadas no proteoma e que também estão no transcriptoma.

TABELA 05: GENES ATIVADOS NO TRANSCRIPTOMA DURANTE A FIXAÇÃO BIOLÓGICA DE NITROGÊNIO CUJAS PROTEÍNAS FORAM IDENTIFICADAS NO PROTEOMA NA MESMA CONDIÇÃO.

Nome	log2 FoldChange	Protein_blast	Vol.
***nif*D**	9,19	**nitrogenase molybdenum-iron protein alpha chain, *nifD***	**0,35**
***nif*K**	8,80	**nitrogenase molybdenum-iron protein beta chain, *nifK***	**0,35**
***nif*U**	8,33	**Fe-S cluster assembly protein**	**0,319**
***nif*H**	8,17	**nitrogenase iron protein, *nifH*, dinitrogenase reductase**	**1,772**
***fix*X**	7,37	**4Fe-4S ferredoxin**	**0,069**
***fix*C**	5,69	**4Fe-4S ferredoxin**	**0,069**

Além dos genes *nif* e *fix* identificados no proteoma em condição de Fixação de Nitrogênio, também foi possível identificar sete proteínas ribossomais e alguns genes do metabolismo energético, como o *suc*D, *ace*B, *atp*CD e *dld* (ver Anexo 04). Os resultados dos proteomas realizados reforçam os resultados dos transcriptomas.

4.2. Re-anotação do Genoma da *Herbaspirillum seropedicae* SmR1

RNA-Seq tem provido informações sobre a expressão gênica de procariotos. Sequenciamento de RNA em diversas condições tem sido utilizados para análises de expressão diferencial de genes, realizar anotação de genomas, anotação de operons e refinamento de anotação gênica (MCCLURE *et al*., 2013; ROBERTS *et al.*, 2011; TJADEN, 2015).

MCCLURE e colaboradores (2013) desenvolveram métodos computacionais para identificar novos transcritos, regiões 5'-UTR, operons, ncRNAs e genes diferencialmente expressos com base em dados de RNA-Seq. Foram utilizados dados de trancriptoma das bactérias *Xenorhabdus nematophila* ATCC 19061, *Neisseria gonorrhoeae* F62, *Streptococcus pyogenes* MGAS2221, *Salmonella enterica* subsp. *enterica serovar Typhimurium* str. LT2 e *Escherichia coli* str. K-12 substr. MG1655 para validar o algoritmo implementado (*software* Rockhopper). A Tabela 06 apresenta a quantidade de novos genes e de operons anotados para cada espécie.

TABELA 06: NOVOS GENES E OPERONS IDENTIFICADOS PARA CINCO ESPÉCIES DE BACTÉRIAS, COM BASE EM DADOS DE RNA-SEQ.

Genoma	Número de condições experimentais	Novos genes (não anti-sense)	Novos genes (antisense)	Operons
Xenorhabdus nematophila ATCC 19061	2	589	198	855
Neisseria gonorrhoeae F62	2	162	20	410
Streptococcus pyogenes MGAS2221	2	227	50	412
Salmonella enterica subsp. *enterica serovar Typhimurium* str. LT2	4	6	-	868
Escherichia coli str. K-12 substr. MG1655	16	107	306	838

Os resultados de MCCLURE e colaboradores (2013) mostram que dados de RNA-Seq podem ser utilizados para elucidar características do genoma que até então estavam desconhecidas; aumentando assim o números de genes conhecidos.

Outro tipo de trabalho que pode ser realizado com dados de RNA-Seq é o ajuste na anotação de códon de início e códon de parada; e anotação de regiões 5'-UTR. *Deinococcus deserti* é uma bactéria do deserto, tolerante à dessecação e à radiação. RNA-seq foi realizado

para explorar as especificidades de seu transcriptoma. O sequenciamento mostrou que para 1.174 (60%) mRNAs, o local de início da transcrição foi encontrado exatamente ou muito próximo do códon de inicio da tradução AUG ou GUG. Genes *leaderless* (sem liderança), que perderam regiões não traduzidas (5'-UTR) em seus mRNAs. Essa proporção de mRNAs *leaderless* que podem se assemelhar a mRNAs ancestrais, é sem precedentes para uma espécie bacteriana. Os dados também permitiram a reanotação da posição do códon de início de 257 genes, incluindo vários genes de reparo de DNA. Além disso, vários novos genes induzidos por radiação foram encontrados, e seus papéis potenciais são estudados para tolerância à radiação, dessecação e adaptação a condições ambientais severas (GROOT *et al*., 2014).

Sequenciamento de RNA de *Saccharopolyspora erythraea* [18], revelou conteúdo substancial de transcrição (32%) originado de sequências não anotadas. A análise dessas regiões revelou transcrição de aproximadamente 300 regiões intergênicas que apresentaram transcrição; e confirmou 164 ORFs anteriormente anotados como proteínas hipotéticas (MARCELLIN, *et al*., 2013).

Eckalbar e colaboradores (2013) realizaram RNA-Seq de 11 tecidos adultos e 3 tecidos embrionários de *Anolis carolinensis* (lagarto verde). Os bancos de dados públicos de Ensembl e NCBI forneceram a primeira anotação do genoma de *Anolis carolinensis,* que se baseia na conservação de seqüências com espécies relacionadas. A segunda anotação baseada em transcriptomas específicos para tecidos fornece mais recursos para estudos moleculares. Esta anotação revisada descreve 59.373 transcrições, em comparação com 16.533 e 18.939 que foram identificadas de acordo com Ensembl (FLICEK, 2012) e NCBI; e 22.962 genes codificadores de proteínas. Uma melhoria importante nesta segunda anotação é a cobertura das sequências da região não traduzida (UTR), com 79% e 59% das transcrições contendo 5' e 3' UTRs, respectivamente, o que possibilita a predição de uma sequência protéica mais precisa e uma análise regulatória. A inclusão de sequência de transcriptoma de tecido específico na anotação do genoma do *A. carolinensis* melhorou sua utilidade para estudos comparativos e funcionais. Esta anotação revisada também fornece um atlas de expressão gênica específica para tecidos adultos e embrionários (ECKALBAR, 2013).

Ademais, utilizando dados de RNA-seq de três condições experimentais de *Herbaspirillum seropedicae* SmR1, Moreno (2013) pôde fazer a predição e confirmar a expressão de 98 ncRNAs (RNA não codificador) nesta bactéria. E Barbosa (2014), utilizando dados de RNA-Seq de outras 8 (oito) condições de cultivo, conseguiu confirmar a expressão de mais 75 ncRNAs. Totalizando 173 ncRNAs com expressão confirmada por RNA-Seq. Alguns

desses ncRNAs possuem expressão em condição específica, como em presença de Nitrato, quando a bateria está aderida a raiz do milho, dentre outras.

Estes exemplos revisados acima traz uma noção da diversidade de características que podem ser anotadas ou aperfeiçoadas ao se explorar as informações contidas no sequenciamento total de RNA.

4.2.1 Novos genes do Genoma da *Herbaspirillum seropedicae* SmR1

A anotação original do genoma da *H. seropedicae* SmR1 foi feita pelo projeto GENOPAR (PEDROSA, 2011) e possui 4804 genes. As bibliotecas de RNA-Seq desse microrganismo, em diferentes condições de cultivo, dão indícios de que pode haver mais genes, pois há transcritos em regiões onde não há genes anotados. Em 2014, utilizando a ferramenta HGF (Hybrid Gene Finder) foram detectados 231 genes ainda não anotados e com evidência de expressão por RNA-Seq, tendo como critério cobertura >=3. E foram identificados também 861 operons, de acordo com o *software* Rockhopper (MCCLURE *et al.*, 2013; BARBOSA, 2014).

O NCBI (2015) anotou 192 genes que não constam na anotação original, sendo também considerados novos genes as mudanças de *frame* e as mudanças de sentido. 33 desses genes possuem expressão em pelo menos uma biblioteca de RNA-Seq, com cobertura >=3. Alguns genes anotados pelo NCBI não tiveram códon de parada respeitado. Essas ocorrências foram corrigidas, pois não foram encontrados indícios, através da análise dos transcritos, de que fosse necessário corrigir algum nucleotídeo no genoma.

As Figuras 16, 17 e 18 mostram exemplos de visualizações feitas utilizando o *software* Artemis com a ferramenta BamView (CARVER, *et al*. 2010), esclarecendo como os dados de RNA-Seq e o GC *frame plot* contribuíram para o refinamento da anotação do genoma.

A Figura 16 a) mostra o gene HSERO_RS06505 que foi anotado pelo NCBI, há evidência de expressão do gene, o GC frame plot também dá indícios de haver um gene na primeira *frame* e há mais de vinte *hits* ao realizar o BLASTp. Portanto o gene HSERO_RS06505 foi adicionado ao genoma. A mesma análise foi feita para os novos genes preditos pelos *softwares* HGF e Glimmer. A Figura 16 b) mostra que o gene Hsero_0645 teve seu tamanho alterado pela anotação do NCBI, sendo o gene HSERO_RS03225. Os transcritos, o GC *frame plot* e o resultado do BLASTp confirmam que o tamanho no gene HSERO_RS03225 está correto. E a Figura 16 c) mostra que o gene Hsero_1901 foi removido da *frame* 01 e adicionado

outro gene na *frame* 02, o HSERO_RS09490. O gráfico GC frame plot confirma a mudança de frame. O BLASTp do gene Hsero_1901 obteve hit apenas com ele mesmo. Já o BLASTp do gene HSERO_RS09490 obteve cinco *hits* com outros genomas e possui domínio funcional.

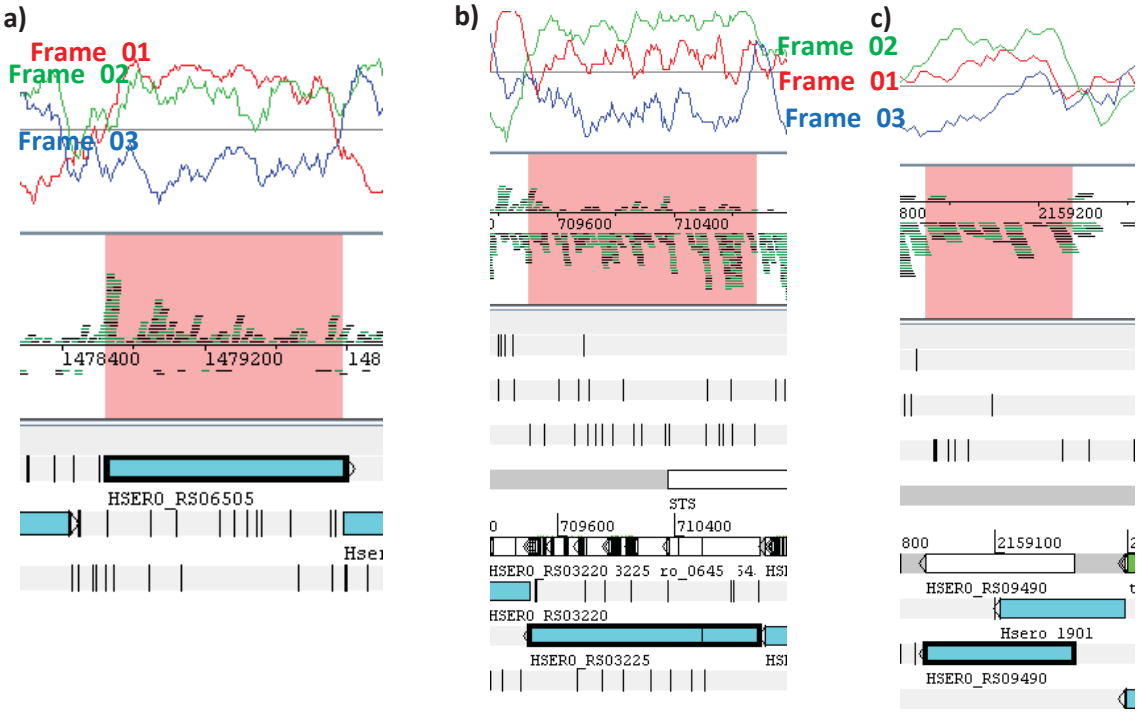


FIGURA 16: EXEMPLOS DE GENES ADICIONADOS AO GENOMA, COM BASE EM DADOS DE RNA-SEQ E NO GRÁFICO GC FRAME PLOT.

a) O gene HSERO_RS06505 anotado pelo NCBI teve sua expressão confirmada por RNA-Seq; b) o gene HSERO_RS03225 anotado pelo NCBI é um ajuste de tamanho do gene Hsero_0645; e c) ajuste de frame do genes Hsero_1901 para o gene HSERO_RS09490.

A Figura 17 a) mostra que o gene Hsero_0372 da anotação original do genoma teve seu tamanho diminuído na anotação do NCBI e adicionou um gene antisense, o HSERO_RS01845. Ao visualizar os transcritos é possível perceber que a anotação Hsero_0372 é a correta. Ao realizar BLASTp é encontrado um domínio funcional no gene Hsero_0372 e no HSERO_RS01845 não. Na Figura 17 b) há um exemplo em que o gene Hsero_2139 da anotação original foi substituído pelo gene HSERO_RS10685 antisense na anotação feita pelo NCBI. Os transcritos estão no sentido do HSERO_RS10685. Ademais, ao realizar o BLASTp do gene Hsero_2139, há *hit* apenas com ele mesmo. E o BLASTp do HSERO_RS10685 há identificação de domínio funcional e o produto 'serine endopeptidase', reforçando a decisão de que é o gene correto.

Na Figura 17 c) o gene Hsero_0854 foi substituído pelo HSERO_RS04290 antisense. Os transcritos estão nos dois sentidos, dificultando decidir qual é o correto. O *frame plot* do gene mostra que o gene deve estar na frame 02. O BLASTp do gene Hsero_0854 retornou *hit* apenas

com ele mesmo, enquanto o BLASTp do gene HSERO_RS04290 retornou mais de 18 hits e encontrou domínio funcional. Portanto foi decidido pelo gene HSERO_RS04290.

A Figura 18 a) mostra o caso do gene Hsero_2587 que foi removido e adicionado dois outros genes na mesma região, os genes HSERO_RS12940 e HSERO_RS12950. Os transcritos confirmam isso e o GC frame plot também. Ao fazer BLASTp do gene Hsero_2587 o resultado dá indícios de que realmente são dois genes, conforme mostra a Figura 18 b).

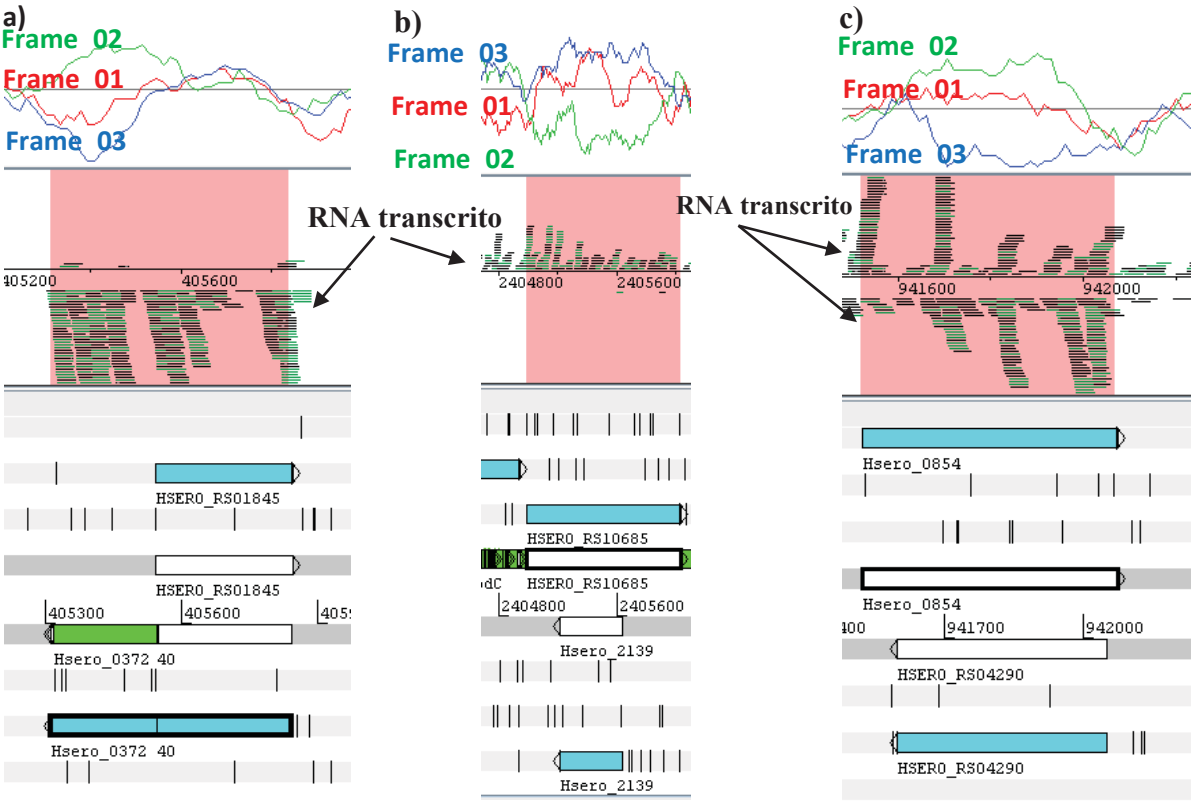


FIGURA 17: EXEMPLOS DE ALTERAÇÕES FEITAS NA ANOTAÇÃO DO GENOMA DA *Herbaspirillum seropedicae* SmR1. a) anotação original mantida; b) gene anotado em outro sentido com base em dados de RNA-Seq; e c) gene anotado em outro sentido com base no GC *frame plot*.

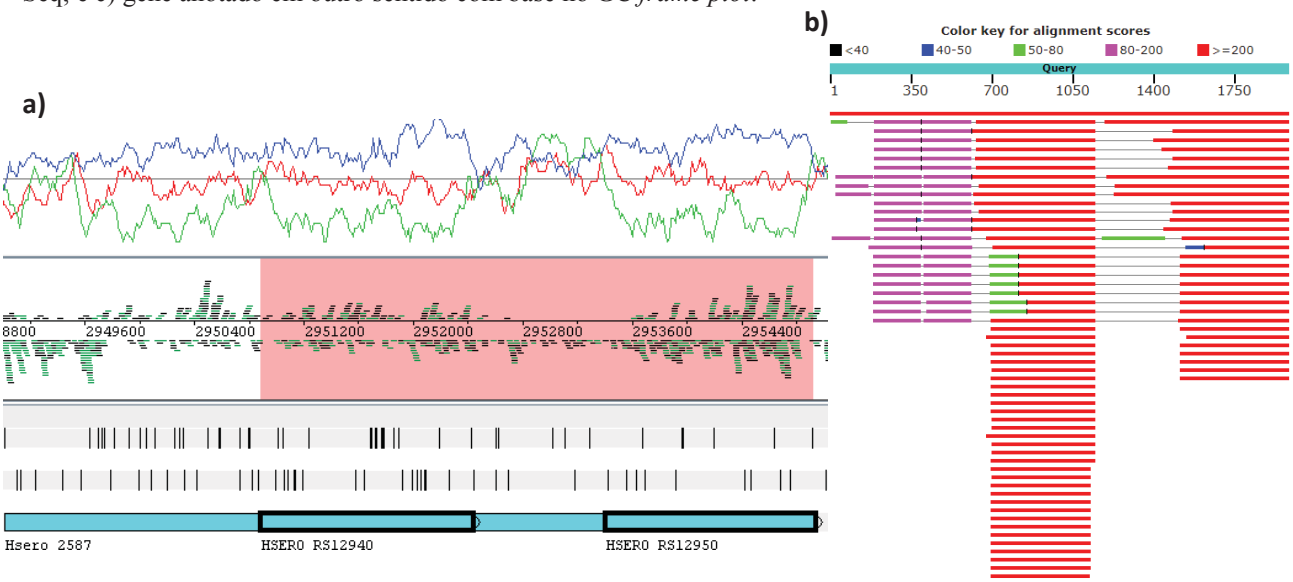


FIGURA 18: a) EXEMPLO DE UM GENE REMOVIDO DO GENOMA, e foi adicionado outros dois na região. b) Resultado do BLASTp para o gene Hsero_2587, indicando que se trata de dois genes.

Neste trabalho também foi utilizado o *software* Glimmer (DELCHER, 1999), que encontrou 123 genes que não possuem equivalência com os já anotados pelo GENOPAR, HGF ou NCBI. Tais genes foram verificados e nenhum possui função conhecida, de acordo com BLASTp e pfam. Quarenta (40) desses genes possuem transcritos observados por RNA-Seq em pelo menos uma condição de cultivo, mas por não possuírem alguma similaridade no banco de dados do NCBI (BLASTp) e o nível de expressão ser muito baixo (alguns 1 *read* por gene), tais genes não foram considerados.

Na continuidade do trabalho de reanotação do genoma da *H. seropedicae* SmR1 as bibliotecas de RNA-Seq, mais as ferramentas BLASTn (nucleotídeos/nucleotídeos), BLASTp (proteínas/proteínas) e BLASTx (nucleotídeos traduzidos/proteínas) e megaBLAST foram utilizadas para verificar a veracidade desses genes (ALTSCHUL *et al.,* 1990).

Depois de compararmos os possíveis novos genes apontados pelo NCBI, HGF e Glimmer montamos uma lista de 309 genes, disponível no ANEXO 05, que podem ser realmente designados como novos, estando portanto ausente na anotação original do Genopar. Para que esses genes fossem incorporados na versão final da re anotação, inicialmente comparamos com o Banco de Dados do NCBI, depois confirmamos novamente se eles estavam expressos em alguma das condições de crescimento analisadas, e a sua posição foi confirmada pelo GC *frame plot.*

Informações contidas no 'GC *frame plot'*, um dos gráficos disponibilizados pelo *software* Artemis, também foram utilizadas para verificar a região mais adequada dos genes. O GC *frame plot* mostra o conteúdo GC da terceira posição do códon para as três *frames* possíveis, sense e antisense. O gene deve estar na *frame* com maior conteúdo GC na terceira posição do códon (ISHIKAWA & HOTTA, 1999).

A Tabela 07, a seguir, mostra quantos genes de cada anotação foi adicionado na versão final da reanotação do genoma. ORFs que não possuíam nenhuma evidência de ser uma região codificadora não foram validadas.

TABELA 07: NÚMERO DE NOVOS GENES ANOTADOS PELO NCBI, PELO HGF E PELO GLIMER RESPECTIVAMENTE, NA REANOTAÇÃO DO GENOMA DA *H. SEROPEDICAE* SMR1.

NCBI	**HGF**	**Glimmer**
192	105	12

Tabela 08 mostra o perfil do genoma da *Herbaspirillum seropedicae* SmR1 após reanotação, comparando com a anotação original. Sendo que os "genes com função desconhecida" compreende também os genes sem classificação funcional. E é possível observar que os genes adicionados possuem uma média de tamanho menor, pois abaixou o tamanho médio das ORFs de 1028,8 para 540,62pb.

TABELA 08: DADOS DA ANOTAÇÃO ORIGINAL (2011) DO GENOMA DE *Herbaspirillum seropedicae* SMR1 E DA RE-ANOTAÇÃO.

	Genopar (2011)	**Reanotação**
Tamanho	5.513.887 bp	5.513.887 bp
Total de genes	4.804	5.037
Total CDS	4735	4973
Tamanho médio das ORFs	1.028,8	540,62
Regiões codificadoras de proteínas	88,3%	90,16%
rRNA operons	3	3
tRNAs	55	55
Operons	-	901
Genes com Função Desconhecida	1.130	751

De acordo com Villegas & Kropinski (2008) 80,1 % dos códons de início em genomas de bactérias são ATG (codifica para Metionina), 11,6 % são GTG (codifica para Valina) e 7,8 % são TTG (codifica para Leucina). Eles examinaram 620 genoma de bactérias disponíveis no GenBank. Na reanotação do genoma da *H. seropedicae* SmR1 90,18% dos códons de início são ATG, 5,44% são GTG e 4,31% são TTG.

Com os dados de expressão gênica contidos nas bibliotecas das onze condições de RNA-Seq foi possível calcular quantos genes no genoma estão expressos em pelo menos uma das condições. O critério para considerar expresso foi cobertura maior ou igual a três. Assim, 88,21% dos genes contidos na reanotação do genoma estão expressos.

4.2.2 Reclassificação funcional do Genoma da *H. seropedicae* SmR1

Com o objetivo de atualizar a classificação funcional dos genes, foi realizado um megablast de todas as proteínas do genoma da *H. seropedicae* SmR1, após a reanotação, com o banco de dados de proteínas do COG, que foi atualizado em 2014 (Clusters of Orthologous Groups - https://www.ncbi.nlm.nih.gov/COG/).

Para realizar a reclassificação COG foi aceito as funções gênicas com no mínimo 30% de similaridade no alinhamento e *e-value* menor igual que 0,001 (PEARSON, 2013).

A tabela 08 mostra o número de genes para cada classe funcional do COG para a anotação inicial, do NCBI e desta reanotação.

Através da tabela 09, podemos observar que categorias que antes não haviam genes contemplados, com a atualização das classes funcionais, agora há genes inclusos, que são os casos das categorias A, B, W, X e Z.

TABELA 09: CLASSIFICAÇÃO FUNCIONAL DOS GENES DO GENOMA DA *H. seropedicae* SmR1 DAS TRÊS VERSÕES DE ANOTAÇÃO.

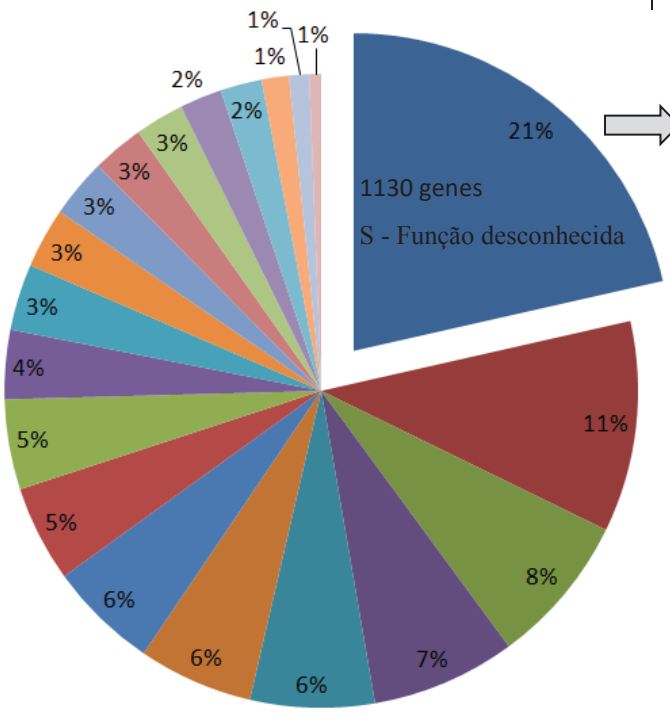
Classificação COG	GENOPAR	NCBI	Re-anotação
A - Processamento e modificação de RNA	0	0	1
B - Estrutura e dinâmica da cromatina	0	0	3
C - Produção e conversão de energia	259	234	261
D - Controle do ciclo celular, divisão celular, partição cromossômica	32	21	47
E - Transporte e metabolismo de aminoácidos	388	438	463
F - Transporte e metabolismo de nucleotídeos	73	67	78
G - Transporte e metabolismo de carboidratos	308	279	292
H - Transporte e metabolismo de coenzima	162	140	218
I - Transporte e metabolismo lipídico	186	158	235
J - Tradução, estrutura ribossomal e biogênese	178	157	246
K - Transcrição	405	330	446
L - Replicação, recombinação e reparação de DNA	113	91	119
M - Parede celular	242	207	323
N - Motilidade celular	158	135	178
O - Modificação pós-traducional, *turnover* de proteína e chaperonas	132	122	155
P - Transporte e metabolismo de íons inorgânicos	294	244	296
Q - Biossíntese, transporte e catabolismo de metabólitos secundários	139	114	155
R - Predição de função geral apenas	567	357	473
S - Função desconhecida	1130	710	181
T - Mecanismos de transdução de sinal	334	268	423
U - Tráfego intracelular, secreção e transporte vesicular	112	100	123
V - Mecanismos de defesa	54	45	100
W - Estruturas extracelulares	0	0	19
X - Elementos móveis: profagos, transposos	0	0	61
Y - Estrutura nuclear	0	0	0
Z - Citoesqueleto	0	0	2

Nessa nova anotação foram identificados dois genes que codificam para proteínas envolvidas com o citoesqueleto. Com a atualização das categorias COG dois genes (Hsero_1591

e Hsero_3846) que estavam na categoria “M - Parede celular” passaram para a categoria “Z - Citoesqueleto”.

Algumas das proteínas citoesqueléticas eucarióticas já são bem conhecidas. Porém, são recentes as pesquisas sobre proteínas citoesqueléticas em procariontes. As bactérias possuem estruturas homólogas ao citoesqueleto que são exclusivos delas (GITAI, 2007). Bactofilin é uma família de uma classe de proteínas citoesqueléticas, formadoras de polímeros, que é amplamente conservada entre as bactérias (KÜHN *et al.,* 2010).

O que há de mais significativo nessa reclassificação é o decaimento do número de genes na categoria “S – Função Desconhecida”. A atualização do Banco de Dados COG possibilitou a inferência de funções gênicas que antes não se conhecia.



Quant.	Classe
181	S - Função desconhecida
80	R - Predição de função geral apenas
49	U - Tráfego intracelular, secreção e transporte vesicular
43	M - Parede celular
36	T - Mecanismos de transdução de sinal
27	K – Transcrição
24	P - Transporte e metabolismo de íons inorgânicos
23	J - Tradução, estrutura ribossomal e biogênese
19	X - Elementos móveis: profagos, transposos
16	O - Modificação pós-traducional, *turnover* de proteína e chaperonas
13	L - Replicação, recombinação e reparação de DNA
12	I - Transporte e metabolismo lipídico
11	E - Transporte e metabolismo de aminoácidos
11	N - Motilidade celular
11	D - Controle do ciclo celular, divisão celular, partição cromossômica
10	C - Produção e conversão de energia
10	Q - Biossíntese, transporte e catabolismo de metabólitos secundários
10	V - Mecanismos de defesa
9	H - Transporte e metabolismo de coenzima
4	W - Estruturas extracelulares
3	G - Transporte e metabolismo de carboidratos
3	F - Transporte e metabolismo de nucleotídeos
1	B - Estrutura e dinâmica da cromatina
570	-Sem Classificação

FIGURA 19: RECLASSIFICAÇÃO DOS GENES “S - FUNÇÃO DESCONHECIDA” NO GENOMA DA *H. seropedicae* SmR1.

Na anotação inicial, a classe funcional mais contemplada é a classe “S - Função desconhecida”, com 1130 genes, o que representa 21,5% dos genes naquela anotação. Na anotação atual a classe “S - Função desconhecida” passa a ter apenas 181 genes.

A figura 19 mostra para quais classes foram os genes “S - Função desconhecida”. As classes mais contempladas são: “R - Predição de função geral apenas”, com 80 genes; seguidas de “U - Tráfego intracelular, secreção e transporte vesicular”, com 49 genes e “M - Parede celular”, com 43 genes.

Ainda restam 570 genes sem classificação. Isso de acordo com os parâmetros de corte que foram utilizados. Ao se ajustar os parâmetros esses valores podem se alterar também. Ficaram 763 genes “Sem Classificação Funcional” por terem ficado fora dos parâmetros de corte. Ser rigoroso na estatística, no momento de inferir função gênica, previne a propagação de erros na função gênica (PEARSON, 2013).

À medida que ocorrem avanços no conhecimento sobre Biologia Molecular, mais se esclarece o papel que as proteínas codificadas pelos genes desempenham nas células. A diminuição do número de genes com “S - Função desconhecida” mostra esse avanço.

A reanotação genômica é fundamental, tanto para identificar todos os genes codificadores de proteínas, como para inferir e atualizar sua função. As técnicas de sequenciamento de nova geração aumentam o número de sequências gênicas conhecidas e estimulam o entendimento do seu papel nas funções biológicas (BERG *et. al,* 2010).

Se um genoma não está anotado corretamente e é submetido a banco de dados públicos outros genomas também podem incorrer em erros, pois a anotação funcional é realizada por comparação de sequências. Assim, o erro pode se propagar. Por isso, é imprescindível atualizar regularmente a anotação genômica de acordo com os bancos de dados também atualizados (GUO *et. al.,* 2013).

Salzberg (2007) propõe que a anotação genômica seja colaborativa. Propõe como solução um novo e expansivo banco de dados que mostre todas as alternativas de anotação de alguma região do genoma, então o pesquisador teria acesso às sobreposições de anotações e aos artigos que contem os experimentos que justificam a mudança de função. Propõe que um "wiki do genoma" possa fornecer a solução que precisamos para a anotação do genoma.

4.2.3 Identificação de operons no Genoma da *H. seropedicae* SmR1

Para identificar os operons de *H. seropedicae* SmR1 foi utilizado o *software* Rockhopper, que localizou 861 operons, com base nos genes anotados juntamente com os dados de RNA-Seq (BARBOSA, 2014). O Rockhopper utiliza a distância entre os genes e a similaridade de expressão, dos genes nos dados de RNA-seq, para determinar a probabilidade de serem transcritos em operon. (MCCLURE *et al.,* 2013).

Dando continuidade a este trabalho, também foi utilizado o banco de dados de operons em procariotos (DOOR - Database of prOkaryotic OpeRons), que utiliza um *software* preditor de operons *online* (MAO, 2009). Neste Banco de dados há a predição de 2785 operons no genoma da *H. seropedicae* Smr1, com a anotação original (PEDROSA, 2011).

Os operons preditos apenas pelo DOOR foram averiguados mediante dados de expressão por RNA-Seq. Os *reads* dos dados de RNA-Seq são soberanos para confirmar, ou não, um operon predito computacionalmente.

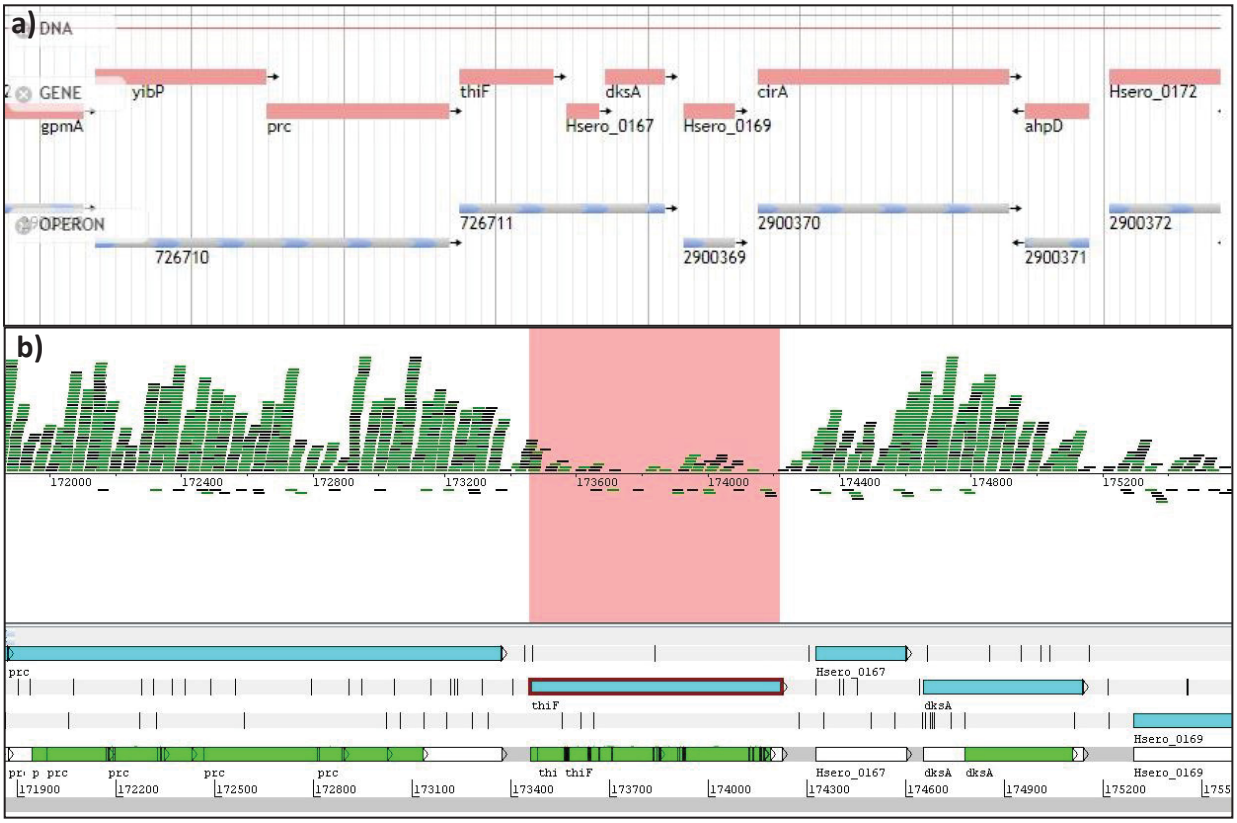


FIGURA 20: a) EXEMPLO DE OPERON COMPOSTO PELOS TRÊS GENES (*thi*F, Hsero_0167 e *dks*A) predito pelo DOOR e **b)** a visualização dos *reads* expressos por RNA-Seq não confirmam tal operon no genoma da *H. seropedicae* SmR1.

A Figura 20 mostra um exemplo de um operon predito e disponível no DOOR, são os genes *thi*F, *Hsero_0167* e *dks*A. Mas ao visualizar os *reads* expressos na biblioteca NFbHP-Malato-Nitrato, tal operon não se confirma. Não há predição desse operon no genoma de algum outro procarioto. Os níveis de expressão entre os três genes são diferentes, indicando que são expressos individualmente.

Com a inclusão de novos genes no genoma, se fez necessário verificar se tais genes participam de operons. A Figura 21 mostra o exemplo de um operon que foi predito como composto por dois genes, *fix*N-*fix*O; e após a inclusão do gene HSERO_RS16045 (*fix*Q), os genes *fix*Q e *fix*P foram adicionados ao mesmo operon; sendo então, um operon composto por quatro genes, *fix*N-*fixO*-*fix*Q-*fix*P. Para incluir novos genes anotados em operons, foi realizado um BLASTn de toda a sequência do candidato a operon, a fim de verificar se são genes vizinhos em outros genomas.

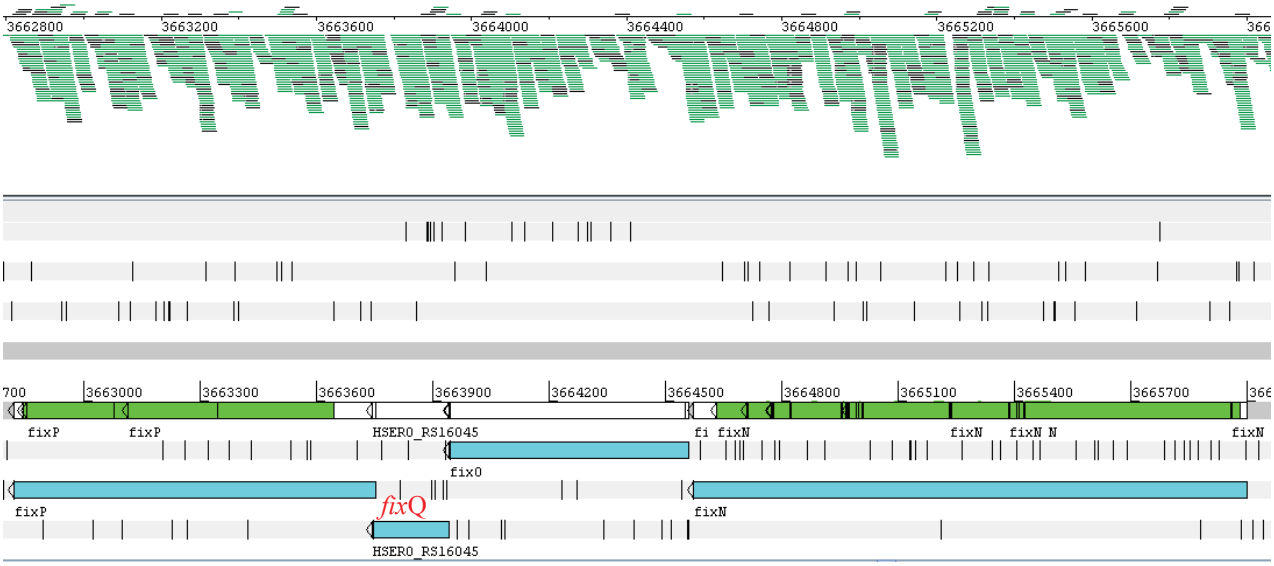


FIGURA 21: EXEMPLO DE AMPLIAÇÃO DE UM OPERON APÓS REANOTAÇÃO DO GENOMA DA *H. seropedicae* SmR1.

O operon livKHMGF na condição NFbHPN-Malato-Nitrato possui os cinco genes no mesmo nível de expressão como mostra na Figura 22 a); enquanto que na condição NFbHPN-Malato-Controle o gene *liv*K está com o nível de expressão maior que os demais genes, Figura 22 b). Trata-se de transportadores de aminoácidos de cadeia ramificada (ADAMS, 1990:).

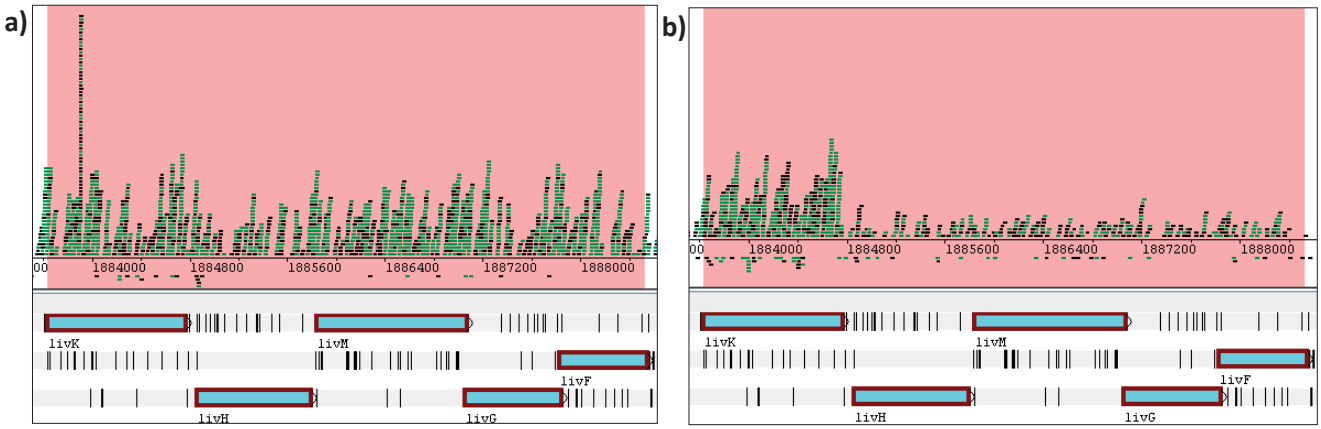


FIGURA 22: OPERON *liv*KHMGF. a) Na condição NFbHPN-Malato-Nitrato e b) na condição NFbHPN-Malato-Controle, onde há o atenuador Lrp.

O mesmo operon ocorre também em *Escherichia coli*, onde foi verificado a ocorrência de dois promotores possíveis para o transcrito livKHMGF. O promotor *liv*Kp1, que ativa o repressor de transcrição Lrp; e o promotor *liv*Kp2, que trancreve todos os genes sem atenuar (HANEY, 1992). Possivelmente, esse evento que acontece em *Escherichia coli* seja o mesmo que também ocorre em *H. seropedicae*. O operon livKHMGF também ocorre em 123 outros genomas de procariotos, de acordo com o banco de dados DOOR.

FIGURA 23: MODELO DO OPERON livKHMGF Em *Escherichia coli,* onde há o promotor *liv*Kp1 com o atenuador Lrp, e o promotor *liv*Kp2.

No operon *glnAntrBC*, *glnA* nem sempre está transcrito junto com *ntrBC*. Há os promotores *glnAp1* e *glnAp2* amontante de glnA; e pntrBC amontante de ntrB. Em condição de cultivo em que o Nitrogênio é suficiente, os promotores *glnAp1* e *pntrBC* são ativos. Em condição em que há privação de Nitrogênio *glnAp1* e *pntrBC* são reprimidos por NtrC e *glnAntrBC* é expresso com o promotor *glnAp2* (PERSUHN *et al.*, 2000).

Em todas as onze condições de transcriptoma por RNA-Seq, utilizadas pelo *software* Rockhopper para predizer operons, há Nitrogênio suficiente. Assim, o Rockhopper não fez a predição do operon *glnAntrBC*, pois há divergência nos níveis de expressão nas condições

analizadas pelo *software*. Mostrando que podem haver outros operons, além dos listados neste trabalho, associados a outras condições de cultivo.

Integrado ao Banco de dados DOOR, há um preditor de terminador, de acordo com o *software* TransTermHP (KINGSFORD *et al*., 2007). Há predição de um terminador ajusante a *glnA* e outro terminador ajusante de *ntrC*. Não há predição do operon *glnAntrBC* pelo DOOR em procariotos. Há a predição do operon *ntrBC* em 248 genomas, mas não em *H. seropedicae.*

De acordo com PERSUHN (2000) o operon *glnAntrBC* foi adicionado a lista de operons preditos em *H. seropedicae*.

O ANEXO 06 contém a relação dos operons do genoma da *H. seropedicae* SmR1 após reanotação.

A tabela 10 mostra a a frequência da quantidade de genes por operon. Há 544 operons compostos por apenas dois genes. E há um operon composto de quinze genes. Um total de 901 operons preditos.

TABELA 10: FREQUÊNCIA DA QUANTIDADE DE GENES POR TAMANHO DE OPERON.

Quantidade de operons	Nº de genes
544	2
193	3
76	4
46	5
23	6
4	7
3	8
3	9
3	10
4	11
1	14
1	15

Os resultados da anotação do genoma com base em dados de RNA-Seq mostram que o transcriptoma fornece informações consistentes para adicionar anotação de genes; ajustar o tamanho, frames e sentido de expressão; e anotar operons.

Conclusões

Em condições de Fixação Biológica de Nitrogênio *Azospirillum brasilense* FP2 ativa o complexo enzimático da Nitrogenase e os genes *nif*, bem como os transportadores de compostos nitrogenados;

Em condições de Fixação Biológica de Nitrogênio *Azospirillum brasilense* FP2 diminui a expressão das vias de metabolismo energético, sendo encontrados reprimidos os genes que participam da Cadeia Respiratória, Ciclo do Ácido Citrico e Glicólise;

Azospirillum brasilense FP2 também teve os genes da Montagem Flagelar reprimidos, o que pode afetar a sua locomoção. Também teve os halos de crescimento medidos, indicando que em condição de Fixação de Nitrogênio há um descrescimo de quatro vezes em sua motilidade;

Ao comparar os resultados de transcriptoma de *Azospirillum brasilense* FP2 com dados de proteoma em condições semelhantes, destaca-se ativados genes *nif* e *fix*;

Ao comparar os resultados de transcriptoma de *Azospirillum brasilense* com dados de trancriptoma de *Azoarcus sp.* BH72, *Azotobacter vinelandii* e *Paenibacilus sp*. WLY78 em condições de Fixação de Nitrogênio é possível destacar em *Azospirillum brasilense* a repressão da Montagem Flagelar.

Os dados de transcriptoma por RNA-Seq possibilitou a validação de 309 novos genes no genoma da *Herbaspirillum seropedicae* SmR1, sendo considerados novos genes os ajustes de *frame* e de sentido;

Também foi possível verificar a expressão de 901 operons;

A reclassificação funcional dos genes de *Herbaspirillum seropedicae* SmR1 permitiu atualizar as classes funcionais e diminuir a quantidade de genes com funções desconhecida.

Referências

ADAMS, M.D.; WAGNER, L. M.; GRADDIS, T. J.; LANDICK, R.; ANTONUCCI, T. K.; GIBSON, A. L.; OXENDER, D. L. Nucleotide sequence and genetic characterization reveal six essential genes for the LIV-I and LS transport systems of *Escherichia coli***. J Biol Chem**; 265(20);11436-43; 1990.

ADLER, S. P.; PURICH, D.; STADTMAN, E. R. Cascade control of Escherichia coli glutamine synthetase. Properties of the PII regulatory protein and the uridylyltransferase-uridylyl-removing enzyme**. J. Biol. Chem**., v.250, n. 16, p. 6264-72, 1975.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. Basic Local Alignment Search Tool. J. Mol. Biol. 215, 403-410. 1990.

ANDERS, S. & HUBER, W. Differential expression analysis for sequence count data. **Genome Biology,** 11, pp. R106, 2010.

ANDERS, S.; REYES, A. and HUBER, W. Detecting differential usage of exons from RNA-seq data. **Genome Research**, 2012.

ARAUJO, M. S.; BAURA, V. A.; SOUZA, E. M.; BENELLI, E. M.; RIGO, L. U.; STEFFENS, M. B.; PEDROSA, F. O.; CHUBATSU, L. S. In vitro uridylylation of the *Azospirillum brasilense* N-signal transducing GlnZ protein. **Protein Expr. Purif.** 33: 19-24, 2004.

ARCONDÉGUY, T; JACK, R; MERRICK, M. PII Signal transduction proteins, pivotal players in microbial nitrogen control. **Microbiology and Molecular Biology Reviews**, v. 65, n. 01, p. 80-105, mar., 2001.

ARNOSTI, D. N. and CHAMBERLIN, M. J. Secondary σ factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*, **Proc. Natl. Acad. Sci. USA.** 86, 830-834, 1989.

ATKINSON, M. R.; KAMBEROV, E. S.; WEISS, R. L.; NINFA, A. J. Reversible uridylylation of the *Escherichia coli* PII signal transduction protein regulates its ability to stimulate the dephosphorylation of the transcription factor nitrogen regulator I (NRI or NtrC**). J. Biol. Chem.** 269: 28288-28293, 1994.

BAGGERLY, K., DENG, L., MORRIS, J., and ALDAZ, C. Differential expression in SAGE: accounting for normal between-library variation. **Bioinformatics**, 19(12):1477-1483. (2003).

BALDANI, V.L.D.; OLIVEIRA, E.; BALOTA, E.; BALDANI, J.I.; KIRCHHOF, G. & DÖBEREINER, J. *Burkholderia brasilensis* sp. nov., uma nova espécie de bactéria diazotrófica endofítica. **An. Acad. Bras. Ci**., 69:116, 1997.

BALSANELLI, E.; **Caracterização de fatores moleculares envolvidos na interação de Herbaspirillum seropedicae com gramíneas.** Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2013.

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

BARBOSA, H. C.S. **Re-Anotação do Genoma de *Herbaspirillum seropedicae* SmR1 com Dados de Transcriptoma por RNA-Seq.** Dissertação de Mestrado, Departamento de Bioinformática, Universidade Federal do Paraná. Curitiba, 2014.

BASHAN, Y.; HOLGUIN, G. and DE-BASHAN, L.E.. *Azospirillum*-plant relationships: Physiological, molecular, agricultural and environmental advances (1997-2003). **Can. J. Microbiol.,** 50: 521-577. 2004.

BATISTA, MARCELO B. ; SFEIR, MICHELLE Z. T. ; FAORO, HELISSON ; WASSEM, ROSELI ; STEFFENS, MARIA B. R. ; PEDROSA, FÁBIO O. ; SOUZA, EMANUEL M. ; DIXON, RAY ; MONTEIRO, ROSE A. . The *Herbaspirillum seropedicae* SmR1 Fnr orthologs controls the cytochrome composition of the electron transport chain. **SCI REP-UK**, v. 3, p. 2544, 2013.

BERG, B. H. J. V. D.; McCARTHY, F. M.; LAMON, S. J.; BURGESS, S. C.; Re-annotation is an essential step in systems biology modeling of functional genomics data. PlosOne. 5(5): e10642. 2010.

BLAHA, D.; PRIGENT-COMBARET, C.; MIRZA, M. S.; MOËNNE-LOCCOZ, Y. Phylogeny of the 1-aminocyclopropane-1-carboxylic acid deaminase-encoding gene *acdS* in phytobeneficial and pathogenic Proteobacteria and relation with strain biogeography. **FEMS Microbiol Ecol,** 56 pp. 455-470. 2006.

BLAIR, D. F. and BERG, H. C. The MotA protein of *E. coli* is a proton-conducting component of the flagellar motor. **Cell**, 60, 439-449, 1990.

BONATO, P. **Genes do metabolismo de Nitrato em *Herbaspirillum seropedicae*: regulação transcricional e análise funcional.** Dissertação (Mestrado em Bioquímica) - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2012.

BUENO, R.; PAHEL, G.; MAGASANIK, B. Role of glnB and glnD gene products in regulation of the glnALG operon of *Escherichia coli*. **J. Bacteriol**. 164: 816-822, 1985.

CARVER, T.; BÖHME, U.; OTTO, T. D.; PARKHILL, J.; BERRIMAN, M. Viewing mapped read alignment data in the context of the reference sequence. **Bioinformatics**, 2010.

CARVER, T.; HARRIS, S. R.; OTTO, T. D.; BERRIMAN, M.; PARKHILL, J.; MCQUILLAN, J. A. Visualizing and interpretation of next-generation sequencing read alignments. **Brief. Bioinform**, 2012.

CHEN, S.; BEEBY, M.; MURPHY, G. E.; LEADBETTER, J. R.; HENDRIXSON, D. R.; BRIEGEL, A.; LI, Z.; SHI, J.; TOCHEVA, E. I.; MULLER, A.; DOBRO, M. J. and JENSEN, G. J. Structural diversity of bacterial flagellar motors. **EMBO J.,** 30, 2972-2981, 2011.

CHILCOTT, G. S. and HUGHES, K. T. Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar *typhimurium* and *Escherichia coli*, **Microbiol. Mol. Biol. Rev.,** 64, 694-708, 2000.

CRUZ, L. M.; SOUZA, E. M.; WEBER, O. B.; BALDANI, J. I.; DÖBEREINER, J.; PEDROSA, F. O. 16S ribosomal DNA characterization of nitrogen-fixing bacteria isolated from banana (*Musa* spp.) and pineapple (*Ananas comosus* (L.) Merril). **Applied and Environmental Microbiology**, v. 67, n.5, p. 2375-2379, 2001.

DAM, P.; OLMAN, V.; HARRIS, K.; SU, Z.; XU, Y. Operon prediction using both genome-specific and general genome information. **Nucleic Acids Research.** 35:288 - 298, 2007.

DE ZAMAROCZY, M. structural homologues PII and PZ of *Azospirillum brasilense* provide intracellular signalling for selective regulation of various nitrogen dependent functions. **Mol. Microbiol.,** Oxford, v. 29(2), p. 449-463, 1998.

DE ZAMAROCZY, M.; PAQUELIN, A.; PELTRE, G.; FORCHHAMMER, K.; ELMERICH, C. Coexistence of two structurally similar but functionally different PII proteins in *Azospirillum brasilense*. **J. Bacteriol**. 178: 4143-4149, 1996.

DELCHER, A. L.; HARMON, D.; KASIF, S.; WHITE, O. and SALZBERG, S. L. Improved microbial gene identification with GLIMMER, **Nucleic Acids Research** 27:23; 4636-4641. 1999.

DIORIO, C.; CAI, J.; MARMOR, J.; SHINDER, R.; DUBOW, M. S. An *Escherichia coli* chromosomal ars operon homolog is functional in arsenic detoxification and is conserved in Gram-negative bacteria. **J. Bacteriol.** 177:2050-2056, 1995.

DIXON, R.; KAHN, D. Genetic regulation of Biological Nitrogen Fixation. **Nature Reviews in Microbiology 2**: 621:631. 2004. 2004

DÖBEREINER, J. History and new perspectives of diazotrophs in association with non-leguminous plants. **Symbiosis** 13: 1–13. 1992.

DÖBEREINER, J.; DAY, J. M. Associative symbioses in tropical grasses characterization of microorganisms and dinitrogen-fixing sites. In: NEWTON, W. E. e NYMAN, C. J. (Ed.) **Proceedings of the 1st. Intl. Symp. Nitrogen Fixation**. v. 2 Pullman, Washington State University, v.2, p.518-538, 1976.

DÖBEREINER, J.; PEDROSA, F. O**. Nitrogen-fixing bacteria in Nonleguminous Crop plants.** New York: Sci.Tech.Publishers/Springer-Verlag 155 p. 1987.

DRUMMOND, M. P.; WHITTY, and WOOTTON, J. Sequence and domain relationships of ntrC and nifA from *Kleshsiella pneumoniae*: homologies to other regulatory proteins. **EMBO J**. 5:441-447. 1986.

DRUMMOND, M.; CLEMENTS, J.; MERRICK, M. & DIXON, R. Positive control and autogenous regulation of the nfLA promoter in *Klebsiella pneumoniae*. *Nature*, London 301, 302-307. 1983.

ECKALBAR, W. L; HUTCHINS, E. D.; MARKOV, G. J.; ALLEN, A. N.; Jason J CORNEVEAUX, J. J.; LINDBLAD-TOH, K.; DI PALMA, F.; ALFÖLDI, J.; HUENTELMAN, M. J. and KUSUMI, K. Genome reannotation of the lizard *Anolis carolinensis* based on 14 adult and embryonic deep transcriptomes. **BMC Genomics**, 2013.

ELMERICH, C.; DE ZAMAROCZY, M.; ARSENE, F.; PEREG, L.; PAQUELIN, A.; KAMINSKI, A. Regulation of *nif* gene expression and nitrogen metabolism in *Azospirillum*. **Soil Biol. Biochem**. 29:847–852, 1997.

FINN, R. D.; MISTRY, J.; TATE, J.; COGGILL, P.; HEGER, A. *et al*. The Pfam protein families database. **Nucleic Acids Res** 38: 211-222, 2009.

FLICEK, P.; AMODE, M. R.; BARRELL, D.; BEAL, K.; BRENT. S.; CARVALHO-SILVA, D.; CLAPHAM, P.; COATES. G.; FAIRLEY, S.; FITZGERALD, S.; GIL, L.; GORDON, L.; HENDRIX, M.; HOURLIER, T.; JOHNSON, N.; KAHARI, A. K.; KEEFE, D.; KEENAN, S.; KINSELLA, R.; KOMOROWSKA, M.; KOSCIELNY, G.; KULESHA, E.; LARSSON, P.; LONGDEN, I.; MCLAREN, W.; MUFFATO, M.; OVERDUIN, B.; PIGNATELLI, M.; PRITCHARD, B.; RIAT, H. S.; RITCHIE, G. R. S.; RUFFIER, M.; SCHUSTER, M.; SOBRAL, D.; TANG, Y. A.; TAYLOR, K.; TREVANION, S.; VANDROVCOVA, J.; WHITE, S.; WILSON, M.; WILDER, S. P.; AKEN, B. L.; BIRNEY, E.; CUNNINGHAM, F.; DUNHAM, I.; DURBIN, R.; FERNANDEZ-SUAREZ, X. M.; HARROW, J.; HERRERO, J.; HUBBARD, T. J. P.; PARKER, A.; PROCTOR, G.; SPUDICH, G.; VOGEL, J.; YATES, A.; ZADISSA, A.; SEARLE, S. M. J. Ensembl 2012. **Nucleic Acids Res**, 40:84–90, 2011.

GAUTHIER, D.; ELMERICH, C. Relationship between glutamine synthetase and nitrogenase in *Spirillum lipoferum*. **FEMS Microbiol**. Lett. 2; 101-104, 1977.

GIEL, J. L.; RODIONOV, D.; LIU, M.; BLATTNER, F. R.; KILEY, P. J. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O_2-regulated genes in *Escherichia coli*. **Mol Microbiol**. 60 (4) :1058-75, 2006.

GITAI, Z. Diversification and specialization of the bacterial cytoskeleton**. Curr Opin Cell Biol.** Feb; 19(1):5-12. 2007.

GRAINGER, D. C.; HURD, D.; GOLDBERG, M. D.; BUSBY, S. J. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. **Nucleic Acids Res.** 34:4642-4652, 2006.

GROOT, A.; ROCHE, D.; FERNANDEZ, B.; LUDANYI, M.; CRUVEILLER, S.; PIGNOL, D.; VALLENET, D.; ARMENGAUD, J. and BLANCHARD, L. RNA Sequencing and Proteogenomics Reveal the Importance of Leaderless mRNAs in the

Radiation-Tolerant Bacterium *Deinococcus deserti*. **Genome Biology and Evolution**, Volume 6, Issue 4, 932–948, 2014.

GUIZELINI, D. **G-FINISHER**: **Uma nova estratégia para refinar e finalizar montagens de genomas bacterianos.** Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2016.

GUO, F. B.; XIONG, L.; TENG, J. L. L.; YUEN, K. Y.; LAU, S. K. P. L.; WOO, P. C.Y. W. Re-annotation of protein-coding genes in 10 complete genomes of *Neisseriaceae* family by combining similarity-based and composition-based methods. DNA Research. Pp. 1-14, 2013.

HAMILTON, T. L.; LUDWIG, M.; DIXON, R.; BOYD, E. S.; SANTOS, P. C. D.; SETUBAL, J. C.; BRYANT, D. A.; DEAN, R, D. and PETERS, J. W. Transcriptional Profiling of Nitrogen Fixation in *Azotobacter vinelandii.* JOURNAL OF BACTERIOLOGY, p. 4477–4486, 2011.

HANEY, S. A.; PLATKO, J. V.; OXENDER, D. L.; CALVO, J. M. Lrp, a leucine-responsive protein, regulates branched-chain amino acid transport genes in *Escherichia coli*. **J Bacteriol**;174(1);108-15; 1992.

HUERGO, L. F.; CHUBATSU, L. S.; SOUZA, E. M.; PEDROSA, F. O.; STEFFENS, M. B.; MERRICK, M. Interactions between PII proteins and the nitrogenase regulatory enzymes DraT and DraG in *Azospirillum brasilense*. **FEBS Lett.** 580: 5232-5236. 2006b.

HUERGO, L. F.; MONTEIRO, R. A.; BONATTO, A. C.; RIGO, L. U.; STEFFENS, M. B. R.; CRUZ, L. M.; CHUBATSU, L. S.; SOUZA, E. M. & PEDROSA, F. O. **Regulation of Nitrogen Fixation in *Azospirillum brasilense*** sp.: cell physiology, plant interactions and agronomic. Fabricio Darío Cassán e Inés García de Salamone (compiladores). Buenos Aires : **Asoc. Argentina de Microbiología**, 2008.

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

IINO, T. Polarity of flagellar growth in *salmonella*, **J. Gen. Microbiol**., 56, 227-239, 1969.

ISHIKAWA, J.; HOTTA, K. FramePlot: a new implementation of the Frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. **FEMS Microbiology Letters.** 174; 251-253; 1999.

JAVELLE, A.; MERRICK, M. Complex formation between AmtB and GlnK: an ancestral role in prokaryotic nitrogen control. **Biochem. Soc. Transact**., v. 33, p. 170−172, 2005.

JIANG, P.; NINFA, A. J. Regulation of autophosphorylation of *Escherichia coli* nitrogen regulator II by the PII signal transduction protein**. J. Bacteriol**. 181: 1906-1911, 1999.

JOERGER, R. D.; JACOBSON, M. R.; PREMAKUMAR, R.; WOLFINGER E. D. and BISHOP, P. E. Nucleotide sequence and mutational analysis of the structural genes (*anf*HDGK) for the second alternative nitrogenase from *Azotobacter vinelandii*. **J. Bacteriol**. 171:1075–1086, 1989.

JONES, C. J. and AIZAWA, S. The bacterial flagellum and flagellar motor: structure, assembly, and function. **Adv. Microb. Physiol**., 32, 109-172, 1991.

JOUANNEAU, Y.; ROBY, C.; MEYER, C. M. & VIGNAIS, P. M. ADP-ribosylation of dinitrogenase reductase in *Rhodobacter capsulatus*. **Biochemistry** 28: 6524-6530. 1989.

KADOWAKI, M. K.; FALARZ, L. J.; ANDRADES, D.; CORDEIRO, F. A.; HUERGO, L. F.; SOUZA, E. M.; PEDROSA, F. O. and CHUBATSU, L. S. Comparative proteomic analysis of *Azospirillum brasilense* wild-type FP2 strain and a constitutive nitrogen fixer mutant. **In: XL Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular – SBBq.** Foz do Iguaçu. Anais do Evento, 2011.

KAMBEROV, E. S.; ATKINSON, M. R.; NINFA, A. J. The *Escherichia coli* PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP**. The Journal of Biological Chemistry,** v. 270, n. 30, p. 17797-18807, 1995.

KANEHISA, M.; GOTO, S.; HATTORI, M.;AOKI-KINOSHITA, K. F.; ITOH, M. *et al.* From genomics to chemical genomics: new developments in KEGG**. Nucleic Acids Res** 34: 354-357, 2006.

KANEHISA, M.; GOTO, S.; SATO, Y.; FURUMICHI, M.; TANABE, M. KEGG for integration and interpretation of large-scale molecular datasets. **Nucleic Acids Res.** 40:D109–D114, 2012.

KANEHISA, M.; GOTO, S.; SATO, Y.; KAWASHIMA, M.; FURUMICHI, M.; TANABE, M. Data, information, knowledge and principle: back to metabolism in KEGG. **Nucleic Acids Research**, vol. 42, p. 199-205, 2013.

KINGSFORD, C.; AYANBULE, K. and SALZBERG, S. L. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. **Genome Biology**. 8:R22, 2007.

KLASSEN, G.; SOUZA, E. M.; YATES, M. G.; RIGO, L .U.; INABA, J.; PEDROSA, F. O. Control of nitrogenase reactivation by the GlnZ protein in *Azospirillum brasilense.* **J. Bacteriol.** 183: 6710-6713. 2001.

KLASSEN, G.; SOUZA, E. M.; YATES, M. G.; RIGO, L. U.; COSTA, R. M.; INABA, J.; PEDROSA, F. O. Nitrogenase switch-off by ammonium ions *in Azospirillum brasilense* requires the GlnB nitrogen signal transducing protein. **Appl. Environ. Microbiol**. 71: 5637-5641. 2005.

KÜHN, J.; BRIEGEL, A.; MÖRSCHEL, E.; KAHNT, J.; LESER, K.; WICK, S.; JENSEN, G. J.; THANBICHLER, M. Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. **The EMBO Journal**. 29, 327–339. 2010.

KURTZ, S.; PHILLIPPY, A.; DELCHER, A. L.; SMOOT, M.; SHUMWAY, M.; ANTONESCU, C. and SALZBERG S. L. Versatile and open software for comparing large genomes. **Genome Biology,** 5:R12, 2004.

KUSTU, S.; SANTERO, E.; KEENER, J.; POPHAM, D. E.; WEISS, D. Expression of σ^{54} (ntrA)-dependent genes is probably united by common mechanism. **Microbiol. Rev.** V.53(3), p. 367-376, 1989.

KUTSUKAKE, K.; OHYA, Y. and IINO, T. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*, **J. Bacteriol.** 172, 741-747, 1990.

LEE, J. W.; HELMANN, J. D. Functional specialization within the Fur family of metalloregulators. **J. D. Biometals**, 20: 485, 2007.

LI, H.; HANDSAKER, B.; WYSOKER, A.; FENNELL, T.; RUAN, J.; HOMER, N.; MARTH, G.; ABECASIS, G.; DURBIN, R. and 1000 Genome Project Data Processing Subgroup - The Sequence alignment/map (SAM) format and SAMtools. **Bioinformatics**, 25, 2078-9, (2009).

LIU, X. and MATSUMURA, P. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. **J. Bacteriol.,** 176, 7345-7351, 1994.

LIU, X.; FUJITA, N.; ISHIHAMA, A. and MATSUMURA, P. The C-terminal region of the alpha subunit of *Escherichia coli* RNA polymerase is required for transcriptional activation of the flagellar level II operons by the FlhD/FlhC complex. **J. Bacteriol.,** 177, 5186-5188, 1995.

LOWERY, R. G.; LUDDEN, P. W. Purification and properties of dinitrogenase reductase ADP-ribosyltransferase from the photosynthetic bacterium *Rhodospirillum rubrum.* **J. Biol. Chem.,** v. 263, p. 16714-16719, 1988.

LOWERY, R. G.; SAARI, L. L.; LUDDEN, P. W. Reversible regulation of the nitrogenase iron protein from *Rhodospirillum rubrum* by ADP-ribosylation in vitro**. J. Bacteriol.** 166:513–518. 1986.

LUDDEN, P. W.; BURRIS, R. H. Activating factor for the iron protein of nitrogenase from *Rhodospirillum rubrum*. **Science**, v. 194, p. 424-426, 1976.

MAO, F.; DAM, P.; CHOU, J.; OLMAN, V.; XU, Y. DOOR: a database for prokaryotic operons. **Nucleic Acids Research**, Vol. 37, p. 459 - 463, 2009.

MAO, X.; MA, Q.; ZHOU, C.; CHEN, X.; ZHANG, H.; YANG, J.; MAO, F.; LAI, W.; and XU, Y. DOOR 2.0: presenting operons and their functions through dynamic and integrated views. **Nucleic Acids Res**; 654-9; 2013.

MARCELLIN, E.; LICONA-CASSANI, C.; MERCER, T. R.; PALFREYMAN, R. W.; and NIELSEN, L. K. Re-annotation of the *Saccharopolyspora erythraea* genome using a systems biology approach. **BMC Genomics,** 14:699, 2013.

MCCLURE, R.; BALASUBRAMANIAN, D.; SUN, Y.; BOBROVSKYY, M.; SUMBY, P.; GENCO, C. A.; ANDERPOOL, C. K.; TJADEN, B.: Computational analysis of bacterial RNA-Seq data. **Nucleic Acids Research**. p. 1-16, 2013.

MERRICK, M. J.; EDWARDS, R. A. Nitrogen control in bacteria. **Microbiol. Rev**., v. 59, p. 604-622, 1995

MIDORIKAWA, T.; MATSUMOTO, K.; NARIKAWA, R.; IKEUCHI, M. An Rrf2-type transcriptional regulator is required for expression of psaAB genes in the *cyanobacterium Synechocystis sp.* PCC 6803**. Plant Physiol.** 151(2):882-92. doi: 10.1104/pp.109.141390, 2009.

MONTEIRO, R. A.; BALSANELLI, E.; WASSEM, R.; MARIN, A. M.; BRUSAMARELLO-SANTOS, L. C. C.; SCHMIDT, M. A.; TADRA-SFEIR, M. Z.; PANKIEVICZ, V. C. S.; CRUZ, L. M.; CHUBATSU, L. S.; PEDROSA, F. O.; SOUZA, E. M. *Herbaspirillum*-plant interactions: microscopical, histological and molecular aspects. **Plant and Soil**, v. xx, p. 01-xx, 2012.

MORENO, L. F. **Predição e análise de expressão de RNAs não codificadores com função regulatória presentes na bactéria** *Herbaspirillum seropedicae* **SmR1.** Dissertação (Mestrado em Bioquímica) – Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2013.

MORETT, E.; BUCK, M. NifA-dependent in vivo protection demonstrates that the upstream activator sequence of nif promoters is a protein binding site. **Proc. Natl. Acad. Sci.** USA, v. 85, p. 9401-9405, 1988.

MORETT, E.; CANNON, W.; BUCK, M. The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of *nif* promoters and can be separated from the positive control function of NifA**. Nucleic Acids Res**. 16:11469–11488. 1988.

MORETT, E.; SEGOVIA, L. The σ54 Bacterial Enhancer Binding-Protein Family: Mechanism of Action and Phylogenetic Relationship of Their Functional Domains. **J. Bacteriol.,** v. 175, p. 6067-6074, 1993.

MORIYA, Y., ITOH, M., OKUDA, S., YOSHIZAWA, A., and KANEHISA, M.; KAAS: an automatic genome annotation and pathway reconstruction server. **Nucleic Acids Res**. 35, W182-W185, 2007.

NINFA, A. J.; ATKINSON, E. S. PII signal transduction proteins. **Trends Microbiol**., v.4, p. 172-9, 2000.

NINFA, A. J.; MAGASANIK, B. Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in Escherichia *coli*. **Proc. Natl. Acad. Sci**. U.S.A., v.83, p.5909-5913. 1986.

NOINDORF, L. Caracterização do operon orf1glnKamtB e do gene amtH de *Herbaspirillum seropedicae*. Curitiba, 138p. Tese (Doutorado em Bioquímica) – Setor de Ciências Biológicas, Universidade Federal do Paraná. 2006.

OHNISHI, K.; KUTSUKAKE, K., SUZUKI, H., and LINO, T. A novel transcriptional regulation mechanism in the flagellar regulon of *Salmonella typhimurium*: an anti-σ factor inhibits the activity of the flagellum-specific σ factor, σ^F, **Mol. Microbiol.** 6, 3149-3157, 1992.

OKON, Y.; ALBRECHT, S. L.; BURRIS, R. H. Carbon and ammonia metabolism of *Spirillum lipoferum*. **J. Bacteriol**. 128: 592, 1976.

OLIVARES, F. L.; BALDANI, V. L. D.; REIS, V. M.; BALDANI, J. I. ; DÖBEREINER, J. Occurrence of the endophytic diazotrophs *Herbaspirillum* spp. in roots, stems and leaves predominantly of Gramineae. **Biol. Fertil. Soils**. 21: 197-200, 1996.

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

OSTERMAN, I. A.; DIKHTYAR, Y. Y.; BOGDANOV, A. A.; DONTSOVA, O. A.; SERGIEV, P. V. Regulation of flagellar gene expression in Bacteria. **Biochemistry (Moscow).** 1447–1456, 2015.

PANKIEVICZ, V. C. S. **Investigação de mecanismos envolvidos na interação entre Herbaspirillum seropedicae e gramíneas.** Projeto de Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2013. (não publicada).

PANKIEVICZ, V. C. S.; CAMILIOS-NETO, D.; BONATO, P.; BALSANELLI, E.; TADRA-SFEIR, M. Z.; FAORO, H.; CHUBATSU, L. S.; DONATTI, L.; WAJNBERG, G.; PASSETTI, F.; MONTEIRO, R. A.; PEDROSA, F. O.; SOUZA, E. M. RNA-seq transcriptional profiling of *Herbaspirillum seropedicae* colonizing wheat (*Triticum aestivum*) roots **Plant Mol Biol** 90: 589. 2016.

PEARSON, W. R. An Introduction to Sequence Similarity ("Homology") Searching. **Curr. Protoc. Bioinformatics**. 2013.

PEDROSA, F. O.; KADOWAKI, M. K.; FALARZ, L. J.; RIGO, L. U.; STEFFENS, M. B. R.; MONTEIRO, R. A.; CRUZ, L. M.; HUERGO, L. F.; SOUZA, E. M.; CHUBATSU, L. S. Differential Proteomic Analysis of *Azospirillum brasilense* FP2 strain Grown under Different Nitrogen Conditions. In: **XXXIX Annual Meeting of The Brazilian Biochemistry and Molecular Biology Society**, Foz do Iguaçu. Anais do Evento, 2010.

PEDROSA, F.O.; MONTEIRO, R.A.; WASSEM, R.; CRUZ, L.M.; *et. al*; Genome of HerbaspirillumseropedicaeStrain SmR1, a Specialized Diazotrophic Endophyte of Tropical Grasses. **PLoS Genetics,** vol. **7,** 2011.

PELANDA, R.; VANONI, M. A.; PEREGO, M.; PIUBELLI, L.; GALIZZI, A.; CURTI, B.; ZANETTI, G. Glutamate synthase genes of the diazotroph Azospirillum brasilense. Cloning, sequencing, and analysis of functional domains. J. Biol. Chem. 268: 3099-3106, 1993.

PERSUHN, D. C.; SOUZA, E. M.; STEFFENS, M. B. R.; PEDROSA, F. O.; YATES, M. G.; RIGO, L. U. The transcriptional activator NtrC controls the expression and activity of glutamine synthetase in *Herbaspirillum seropedicae.* FEMS Microbiology Letters. 192; 217-221; 2000.

PIRO, V. C.; FAORO, H.; WEISS, V. A.; STEFFENS, M. B.; PEDROSA, F. O.; SOUZA, E. M.; and RAITTZ, R. T. FGAP: an automated gap closing tool. **BMC Research Notes,** 7(1):371, 2014.

POPE, M. R.; MURREL, S. A.; LUDDEN, P. W. Covalent modification of the iron protein of nitrogenase from *Rhodospirillum rubrum* by adenosine diphosphoribosylation by specific arginyl residue. **Proceedings of the National Academy of Sciences**, Washington, v. 82, p. 3173-3177, 1985.

POSTGATE, J. R. The Fundamentals of Nitrogen Fixation. Cambridge University Press: **Cambridge**, UK. 1982.

RAITTZ, R. T. A new hybrid bioinformatics platform for the identification of proteins coding regions in prokaryotes. **Abstracts Book: 1st International Workshop on Bioinformatics**. 2010. Disponível em: www.bioinfo.ufpr.br/iwb

RAMOS, R.T.J.; CARNEIRO, A. R.; BAUMBACH, J; AZEVEDO, V.; SCHNEIDER, P. C.; SILVA, A. Analysis of quality raw data of second generation sequences with Quality Assessment Software. **BMC Research Notes,** 4:130, 2011.

ROBERTS, A.; PIMENTEL, H.; TRAPNELL, C. and PACHTER, L. Identification of novel transcripts in annotated genomes using RNA-Seq. **Bioinformatics**. Vol. 27, 2325–2329, 2011.

ROBINSON, M. D. and OSHLACK, A. A scaling normalization method for differential expression analysis of RNA-seq data. **Genome Biology** 11, R25. 2010.

RODIONOV, D. A.; DUBCHAK, I. L.; ARKIN, A. P.; ALM, E. J.; GELFAND, M. S. Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. **PLoS Comput Biol.** 1(5):e55, 2005.

RODIONOV, D. A.; GELFAND, M. S.; TODD, J. D.; CURSON, A. R.; JOHNSTON, A. W. Computational reconstruction of iron- and manganese-responsive transcriptional networks in alpha-proteobacteria. **PLoS Comput Biol**. 15; 2(12):e163, 2006.

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

ROSS, W.; THOMPSON, J. F.; NEWLANDS, J. T.; GOURSE, R. L. *E.coli* Fis protein activates ribosomal RNA transcription in vitro and in vivo. **EMBO J.** 9:3733-3742, 1990.

RUMBLE, S. M.; LACROUTE, P.; DALCA, A.V., FIUME, M., SIDOW, A., et al. SHRiMP: Accurate Mapping of Short Color-space Reads. **PLoS Comput Biol 5(5),** 2009.

RUTHERFORD, K; PARKHILL, J; CROOK, J; HORSNELL, T; RICE, P; RAJANDREAM, M. A.; BARRELL, B. Artemis: sequence visualization and annotation. **Bioinformatics** (Oxford, England) v. 16; nº 10; p. 944-5. 2000.

SAARI, L. L.; POPE, M.; MURRELL, S. *et al.* Studies on the activating enzyme for iron protein of nitrogenase from *Rhodospirillum rubrum*. **J Biol Chem** 261:4973–4977. 1986.

SAARI, L. L.; TRIPLETT, E. W.; LUDDEN, P. W. Purification and properties of the activating enzyme for iron protein of nitrogenase from the photosynthetic bacterium *Rhodospirillum rubrum*. **J Biol Chem** 259:15502-15508. 1984.

SALZBERG, S. L. Genome re-annotation: a wiki solution? **Genome Biology**. 8:102, 2007.

SALZBERG, S.; DELCHER, A.; KASIF, S. and WHITE, O. Microbial gene identification using interpolated Markov models. **Nucleic Acids Research** 26:2, 544-548, 1998.

SARKAR, A.; REINHOLD-HUREK, B. Transcriptional Profiling of Nitrogen Fixation and the Role of NifA in the Diazotrophic Endophyte *Azoarcus* sp. Strain BH72. **PLOS ONE**, 9(2): e86527, 2014.

SHI, H.; WANG, L.; LI, X.; LIU, X.; HAO, T.; HE, X. and CHEN, S. Genome-wide transcriptome profiling of nitrogen fixation in *Paenibacillus sp*. WLY78**. BMC Microbiology**, 2016.

SIMPSON, F. B.; BURRIS, R. H. A nitrogen pressure of 50 atmospheres does not prevent evolution of hidrogen by nitrogenase. **Science**, v. 224, p. 1095-1096, 1984.

SORENSON, M. K.; RAY, S. S. and DARST, S. A. Crystal structure of the flagellar σ/anti-σ complex σ^{28}/FlgM reveals an intact σfactor in an inactive conformation. **Mol. Cell,** 14, 127-138, 2004.

SOUZA, E. M.; FUNAYAMA, S.; RIGO, L.U.; YATES, M. G.; PEDROSA, F. O. Sequence and structural organization of a nifA-like gene and part of a nifB-like gene of *Herbaspirillum seropedicae* strain Z78. **Canadian Journal of Microbiology**, Ottawa, v.137, p.1511-1522, 1991.

SPAEPEN, S.; JOS, V.; ROSELINE, R. Indole-3-acetic acid in microbial and microorganism-plant Signaling. **FEMS Microbiol Rev**. ;31:1–24. doi: 10.1111/j.1574-6976.2006.00058.x. 2007.

STRAGIER, P.; PATTE, J. C. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli.* III. Nucleotide sequence and regulation of the lysR gene. **J. Mol. Biol.** 168:333-350, 1983.

TADRA-SFEIR, M.Z. **Caracterização funcional de genes de Herbaspirillum seropedicae regulados pelo flavonóide naringenina**. Tese de doutorado,

Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2011.

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

TATUSOV, R. L.; FEDOROVA, N. D.; JACKSON, J. D.; JACOBS, A. R.; KIRYUTIN, B. *et al*. The COG database: an updated version includes eukaryotes**.** **BMC Bioinformatics**, 2003.

TJADEN, B. De novo assembly of bacterial transcriptomes from RNA-seq data. **Genome Biology.** 2015.

VAN DOMMELEN, A.; KEIJERS, V.; SOMERS, E.; VANDERLEYDEN, J. Cloning and characterization of the *Azospirillum brasilense* glnD gene and analysis of a glnDmutant. **Mol. Genet. Genomics**, 266: 813-820, 2002.

VIALLE, R. A. **SILA – Um Sistema para Anotação Automática de Genomas Utilizando Técnicas Independente de Alinhamento.** Dissertação – Departamento de Bioinformática – Universidade Federal do Paraná. Curitiba. 2013.

VIALLE, R. A.; PEDROSA, F. O.; WEISS, V. A.; GUIZELINI, D.; TIBAES, J. H.; MARCHAUKOSKI, J. N.; SOUZA, E. M.; RAITTZ, R. T. RAFTS3: Rapid Alignment-Free Tool dor Sequence Similarity Search. **BioRxiv**, 2016.

VILLEGAS, A; KROPINSKI, A. M. An analysis of initiation codon utilization in the Domain Bacteria – concerns about the quality of bacterial genome annotation. **Microbiology**; 154; 2559–2561; 2008.

WANG, S.; FLEMING, R. T.; WESTBROOK, E. M.; MATSUMURA, P., and MCKAY, D. B. Structure of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of transcription, **J. Mol. Biol.,** 355, 798-808, 2006.

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

WOLD, S.; CROOKE, E.; SKARSTAD, K. The *Escherichia coli* Fis protein prevents initiation of DNA replication from oriC *in vitro*. **Nucleic Acids Res.** 24:3527-3532, 1996.

ZASLAVER, A.; MAYO, A. E.; ROSENBERG, R.; BASHKIN, P.; SBERRO, H.; TSALYUK, M.; SURETTE, M. G. and ALON, U. Just-in-time transcription program in metabolic pathways. **Nat. Genet.,** 36, 486-491, 2004.

ZHANG, Y.; BURRIS, R. H.; LUDDEN, P. W.; ROBERTS, G. P. Regulation of nitrogen fixation in *Azospirillum brasilense*. **FEMS Microbiol. Rev**. v. 152, p195-204, 1997.

ZUMFT, W. G. Regulation of nitrogenase activity in anoxygenic phototrophic bacteria. In: EVANS, H. J.; BOTTOMLEY, P. J.; NEWTON, W. E., (Ed.). **Nitrogen fixation research progress.** Dordrecht: Martinus Nijhoff, 1985. p. 551-557.

Anexo 01: Lista de genes reprimidos da *A. brasilense* FP2 em condição de Fixação Biológica de Nitrogênio.

ID	KO	Fold Change CLC	P-value CLC	Fold Change DESeq	P-value DESeq	Fold_change EdgeR	P-value EdgeR	Function
c1.0001.0002	0	-4,5366293	1E-06	-2,3566823	0,00034	0,39158981	5E-05	c1.0001.0002 Hypothetical Protein 1275:1565 forward
c1.0001.0021	0	-6,1057413	2E-07	-3,2906945	0,01272	0,29632405	0,001	c1.0001.0021 pirin 17429:18361 forward
c1.0001.0023	0	-8,5019353	0,0027	-4,4229749	7,9E-05	0,22085475	9E-11	c1.0001.0023 ABC transporter 19092:20204 reverse
c1.0001.0027	K01867	-4,8406432	8E-05	-2,5203	0,00114	0,37235843	2E-05	c1.0001.0027 tryptophanyl-tRNA synthetase 24730:25728 forward
c1.0001.0042	K16165	-12,21538	7E-11	-6,4530788	3,4E-05	0,1468436	3E-07	c1.0001.0042 fumarylacetoacetate hydrolase 37825:38508 forward
c1.0001.0071	0	-12,117967	2E-22	-6,4755724	9,2E-10	0,14694849	7E-15	c1.0001.0071 hypothetical protein 69989:70528 forward
c1.0001.0072	K01869	-5,5275579	5E-06	-2,9152661	1,8E-06	0,32807595	1E-09	c1.0001.0072 leucyl-tRNA synthetase 70640:73210 forward
c1.0001.0074	K02340	-4,6542981	1E-05	-2,5372829	0,04593	0,37412383	0,014	c1.0001.0074 DNA polymerase III subunit delta 73798:74820 forward
c1.0001.0098	K03628	-5,9847743	0,0005	-3,1089071	2,2E-06	0,30600113	9E-09	c1.0001.0098 transcription termination factor Rho 95859:97115 forward
c1.0001.0099	K03071	-9,0107393	4E-09	-4,7448667	2,4E-11	0,1993447	4E-16	c1.0001.0099 preprotein translocase subunit SecB 97258:97755 reverse
c1.0001.0104	K02503	-7,8064112	2E-10	-4,1724745	0,00161	0,23440658	5E-05	c1.0001.0104 hypothetical protein 101666:102034 reverse
c1.0001.0108	0	-13,382895	2E-06	-7,2504875	1,3E-05	0,14594445	4E-08	c1.0001.0108 acetyltransferase 103958:104500 reverse
c1.0001.0114	K03667	-7,7374973	0,0052	-3,9010855	0,00179	0,23406454	4E-05	c1.0001.0114 ATP-dependent protease 108765:109520 forward
c1.0001.0143	0	-8,7143494	0,002	-4,5926114	0,01728	0,22581814	0,004	c1.0001.0143 peptidase 136967:137644 forward
c1.0001.0153	K03820	-6,9265369	2E-06	-3,8178749	0,00281	0,26845589	1E-04	c1.0001.0153 acyltransferase 145285:146874 forward
c1.0001.0155	K00789	-4,10596	2E-20	-2,2406113	0,00045	0,43092088	3E-05	c1.0001.0155 S-adenosylmethionine synthetase 147732:148901 forward
c1.0001.0159	K02600	-5,3114887	3E-10	-2,8219425	1,1E-06	0,33782921	4E-11	c1.0001.0159 transcription elongation factor NusA 150847:152382 forward
c1.0001.0160	K07742	-6,5127495	0,0003	-3,3547808	4,5E-05	0,27730719	7E-07	c1.0001.0160 hypothetical protein 152413:153132 forward
c1.0001.0164	K02956	-4,6130549	7E-26	-2,4784749	2,9E-05	0,38477489	2E-09	c1.0001.0164 30S ribosomal protein S15 157519:157788 forward
c1.0001.0165	K00962	-5,9827599	2E-09	-3,1759256	2,2E-07	0,30027164	3E-14	c1.0001.0165 polynucleotide phosphorylase 158085:160205 forward
c1.0001.0174	0	#DIV/0!	5E-06	#DIV/0!	0,00662	0,01348432	6E-04	c1.0001.0174 hypothetical protein 167301:167894 forward
c1.0001.0180	K01507	-5,4303605	0,0011	-2,8082545	0,0001	0,33751871	9E-07	c1.0001.0180 inorganic pyrophosphatase 174362:174919 reverse
c1.0001.0188	0	-5,0840968	2E-07	-2,6639294	0,02699	0,34754043	0,005	c1.0001.0188 conserved exported protein of unknown function 185130:185591 forward

c1.0001.0190	K02968	-12,051213	1E-82	-6,5841802	4,2E-09	0,1472361	6E-13	c1.0001.0190 30S ribosomal protein S20 186621:186881 forward
c1.0001.0191	K02313	-5,283671	0,0005	-2,7785691	3E-05	0,34906301	4E-07	c1.0001.0191 chromosomal replication initiation protein 187501:189045 forward
c1.0001.0199	0	-6,9288633	4E-42	-3,8148497	8,6E-05	0,25620104	1E-06	c1.0001.0199 hypothetical protein ABAZ39
c1.0001.0210	0	-6,0512853	0,0125	-3,1204572	0,00973	0,31270436	5E-06	c1.0001.0210 integral membrane protein linked to a cation pump 207325:207849 forward
c1.0001.0214	K00029	-4,0798655	0,0002	-2,1423714	0,00402	0,44691003	8E-05	c1.0001.0214 NADP-dependent malic enzyme 211546:213816 forward
c1.0001.0226	K01738	-6,7691766	1E-05	-3,5055586	2,4E-07	0,26476963	3E-09	c1.0001.0226 O-acetyl serine sulfhydrylase A 224982:225941 reverse
c1.0001.0254	K05524	-5,7131781	3E-12	-3,001323	0,01302	0,30251849	0,003	c1.0001.0254 4Fe-4S ferredoxin 255122:255454 forward
c1.0001.0263	K02916	-7,6595234	2E-45	-4,0636465	1,3E-05	0,22326665	3E-07	c1.0001.0263 50S ribosomal protein L35 264380:264580 forward
c1.0001.0265	K01889	-7,2165893	1E-10	-3,8320962	2,3E-05	0,24779652	5E-08	c1.0001.0265 phenylalanyl-tRNA synthetase subunit alpha 265096:266163 forward
c1.0001.0266	K01890	-8,7772858	8E-09	-4,7204479	1E-05	0,20835082	1E-08	c1.0001.0266 phenylalanyl-tRNA synthetase subunit beta 266177:268573 forward
c1.0001.0272	0	-6,2051468	3E-05	-3,2133729	0,00106	0,28718151	7E-05	c1.0001.0272 membrane protein 273692:274444 reverse
c1.0001.0273	K03596	-4,9507594	1E-07	-2,6353045	5,1E-05	0,36516667	2E-07	c1.0001.0273 elongation factor 4 274609:276411 reverse
c1.0001.0282	K01424	-3,7268497	2E-07	-2,0314042	0,04975	0,47121707	0,011	c1.0001.0282 L-asparaginase II 285045:286145 reverse
c1.0001.0283	0	-7,2364574	2E-08	-3,9082134	0,01545	0,24938537	0,003	c1.0001.0283 rRNA methyltransferase 286274:286858 forward
c1.0001.0287	K06941	-4,5395199	6E-05	-2,3938441	0,00178	0,40315562	2E-05	c1.0001.0287 50S rRNA methyltransferase 289203:290372 forward
c1.0001.0290	K01940	-6,6185092	0,0002	-3,448196	3E-06	0,27648159	4E-09	c1.0001.0290 argininosuccinate synthase 291067:292302 forward
c1.0001.0295	K15270	-8,507902	0,0247	-4,1907475	0,03053	0,21486573	0,005	c1.0001.0295 membrane protein 295717:296568 reverse
c1.0001.0296	K02200	-13,410443	3E-08	-7,1655513	5,4E-05	0,13680583	7E-07	c1.0001.0296 cytochrome C biogenesis protein 296772:298220 reverse
c1.0001.0298	K02199	-17,337918	1E-15	-9,2106624	4,2E-05	0,09947171	1E-06	c1.0001.0298 thiol oxidoreductase DsbE 298705:299241 reverse
c1.0001.0306	K12410	#DIV/0!	5E-05	#DIV/0!	0,00964	0,01451666	9E-04	c1.0001.0306 NAD-dependent deacetylase 304799:305530 forward
c1.0001.0312	K01491	-6,8986406	0,0011	-3,6218861	0,01838	0,27583013	0,002	c1.0001.0312 5,10-methylene-tetrahydrofolate dehydrogenase 308863:309771 reverse
c1.0001.0321	K07085	-5,1552061	0,0236	-2,6151027	0,02914	0,36289784	6E-04	c1.0001.0321 transporter 315343:316992 reverse
c1.0001.0322	K07085	-4,9031086	2E-06	-2,6714007	0,00893	0,36175084	7E-04	c1.0001.0322 transporter 317094:318773 reverse
c1.0001.0341	K01696	-7,3356988	3E-10	-3,9729906	1,8E-05	0,25143932	3E-08	c1.0001.0341 tryptophan synthase subunit beta 344969:346186 forward
c1.0001.0342	K01695	-6,2934931	0,0015	-3,2463861	0,00037	0,29192517	9E-06	c1.0001.0342 RPA
c1.0001.0343	K01963	-9,0506811	2E-08	-4,8111581	4,3E-08	0,20160552	8E-13	c1.0001.0343 acetyl-CoA carboxyl transferase 347101:348081 forward

c1.0001.0361	0	-4,0075413	4E-14	-2,1210861	0,00041	0,44156894	5E-05	c1.0001.0361 Hypothetical Protein 365396:366511 reverse
c1.0001.0362	0	-5,2005149	2E-08	-2,7034482	0,00011	0,33757691	1E-05	c1.0001.0362 Hypothetical Protein 366656:366940 forward
c1.0001.0387	K06915	-7,7655374	1E-07	-4,0671864	1,8E-09	0,2306506	8E-13	c1.0001.0387 nucleoside triphosphate hydrolase 398397:400082 reverse
c1.0001.0389	0	-6,3760023	0,0034	-3,2084539	0,04918	0,27877482	0,023	c1.0001.0389 Hypothetical Protein 401240:402058 reverse
c1.0001.0390	0	-6,0789165	0,0232	-3,110711	0,01879	0,31020504	6E-07	c1.0001.0390 Hypothetical Protein 402055:406293 reverse
c1.0001.0394	0	-4,8559714	8E-10	-2,5844362	0,00107	0,36784292	9E-06	c1.0001.0394 Hypothetical Protein 408543:409715 reverse
c1.0001.0436	K09780	-4,0968121	2E-30	-2,1993098	0,00675	0,4281269	4E-04	c1.0001.0436 hypothetical protein ABAZ39
c1.0001.0442	0	-12,01907	9E-05	-6,2298482	0,01774	0,15827855	0,004	c1.0001.0442 hypothetical protein 459005:459202 reverse
c1.0001.0444	0	-5,5809967	0,0044	-2,9087504	0,00243	0,33657681	4E-06	c1.0001.0444 exopolyphosphatase 460586:461518 forward
c1.0001.0457	0	-10,158528	0,0122	-5,2020928	0,00273	0,18965107	2E-04	c1.0001.0457 methyl-accepting chemotaxis protein 480265:481962 reverse
c1.0001.0461	K19055	-6,9129248	2E-05	-3,577198	0,01823	0,25831573	0,004	c1.0001.0461 DNA-binding protein 486229:486795 forward
c1.0001.0466	0	-5,2768491	1E-10	-2,856154	0,04132	0,33549895	0,011	c1.0001.0466 hypothetical protein ABAZ39
c1.0001.0473	K07552	-17,80327	0,003	-9,6559838	0,0493	0,12805695	0,006	c1.0001.0473 MFS transporter permease 493859:495043 reverse
c1.0001.0481	K00799	-43,781169	0,0007	-23,740226	0,00015	0,05252703	7E-06	c1.0001.0481 glutathione S-transferase 502627:503232 forward
c1.0001.0485	K11927	-4,8459612	2E-10	-2,6002059	0,00024	0,37241404	1E-06	c1.0001.0485 DEAD/DEAH box helicase 505621:507057 reverse
c1.0001.0512	K02342	-5,2985743	0,0064	-2,7172609	0,0021	0,34868061	2E-06	c1.0001.0512 exonuclease 536114:538321 reverse
c1.0001.0519	K03110	-5,2018525	2E-08	-2,7961286	0,00594	0,34655682	4E-04	c1.0001.0519 cell division protein FtsY 544651:545910 reverse
c1.0001.0520	K18707	-4,3988548	0,008	-2,2552886	0,00841	0,41911242	8E-04	c1.0001.0520 2-methylthioadenine synthase 545927:547186 reverse
c1.0001.0526	K01704	-8,7036118	4E-45	-4,6920806	7,7E-08	0,20252433	1E-11	c1.0001.0526 isopropylmalate isomerase 552045:552656 forward
c1.0001.0527	K00052	-8,9404184	2E-87	-4,7987193	3,2E-13	0,19622489	3E-17	c1.0001.0527 3-isopropylmalate dehydrogenase 552780:553895 forward
c1.0001.0530	K01644	-20,905621	0,0001	-11,561096	0,02162	0,10673232	0,004	c1.0001.0530 malyl-CoA thiolesterase 555082:555963 reverse
c1.0001.0531	K00868	-7,7736257	0,0005	-4,0108477	1,3E-06	0,23434828	2E-09	c1.0001.0531 pyridoxamine kinase 556131:556991 forward
c1.0001.0532	K06966	-6,2683209	6E-16	-3,4671049	0,00133	0,2860985	7E-05	c1.0001.0532 LOG family protein 557040:557618 forward
c1.0001.0533	0	-6,9370663	2E-10	-3,7162003	0,0218	0,262686	0,004	c1.0001.0533 conserved exported protein of unknown function 557631:558044 reverse
c1.0001.0534	K06916	-11,527876	3E-07	-6,076036	8,3E-08	0,15788379	1E-11	c1.0001.0534 ATPase 558178:559332 reverse
c1.0001.0537	K00658	-11,081797	2E-31	-5,900785	4,3E-15	0,15896913	5E-20	c1.0001.0537 dihydrolipoamide succinyltransferase 563448:564701 reverse
c1.0001.0539	K01902	-4,1475139	0,0037	-2,1365326	0,00318	0,44224446	2E-04	c1.0001.0539 succinyl-CoA synthetase subsunit alpha 567818:568693 reverse
c1.0001.0540	K01903	-7,2290699	5E-06	-3,824033	9,7E-10	0,25152399	5E-14	c1.0001.0540 malate--CoA ligase subunit beta 568695:569891 reverse

c1.0001.0541	K00024	-10,577374	4E-09	-5,5816533	3,9E-15	0,17036697	1E-23	c1.0001.0541 malate dehydrogenase 570144:571112 reverse
c1.0001.0552	K00651	-8,4011745	0,0028	-4,2000597	0,00355	0,20961994	0,001	c1.0001.0552 homoserine O-succinyltransferase 583253:584200 reverse
c1.0001.0555	K07088	-13,549573	8E-11	-7,5939961	4,6E-05	0,14624557	2E-07	c1.0001.0555 transporter 586944:587891 reverse
c1.0001.0556	0	-5,4525477	8E-14	-2,9065715	0,00025	0,32748151	9E-07	c1.0001.0556 hypothetical protein 588039:588443 forward
c1.0001.0561	K00241	-4,5869593	7E-16	-2,4240075	0,00703	0,37563876	0,001	c1.0001.0561 succinate dehydrogenase 591467:591856 forward
c1.0001.0562	K00242	-4,6400828	2E-46	-2,5069524	0,00026	0,38012189	2E-06	c1.0001.0562 succinate dehydrogenase 591868:592263 forward
c1.0001.0563	K00239	-5,4093326	1E-09	-2,8773732	1E-06	0,33230766	2E-10	c1.0001.0563 fumarate reductase 592270:594051 forward
c1.0001.0564	K00240	-7,3063037	0,0318	-3,723825	0,02411	0,25892763	2E-08	c1.0001.0564 succinate dehydrogenase iron-sulfur subunit 594082:594858 forward
c1.0001.0583	0	-66,946789	4E-07	-34,912698	2,6E-42	0,02696948	3E-66	c1.0001.0583 Hypothetical Protein 605083:606498 forward
c1.0001.0584	0	-67,94744	0,0001	-35,028003	5,4E-19	0,02704197	2E-35	c1.0001.0584 Hypothetical Protein 606495:606830 forward
c1.0001.0585	K16263	-38,701825	5E-05	-19,962142	1,2E-21	0,04682879	8E-38	c1.0001.0585 Hypothetical Protein 606817:608124 forward
c1.0001.0586	0	-112,34389	0,0001	-55,536934	2E-18	0,01518901	5E-31	c1.0001.0586 Hypothetical Protein 608206:608655 forward
c1.0001.0587	0	-27,503331	6E-43	-15,236714	4,4E-10	0,06861107	3E-14	c1.0001.0587 Hypothetical Protein 608837:609262 forward
c1.0001.0651	0	-8,4520117	5E-05	-4,3526244	0,02436	0,21110291	0,005	c1.0001.0651 HAD family hydrolase 670937:671812 reverse
c1.0001.0652	0	-13,68319	5E-08	-7,2498824	0,00015	0,13181891	4E-06	c1.0001.0652 glycosyl transferase 671851:673092 reverse
c1.0001.0683	0	-6,7230947	6E-15	-3,6075145	0,00305	0,26448257	2E-04	c1.0001.0683 exported protein of unknown function 704425:704916 reverse
c1.0001.0684	0	-8,2935509	6E-50	-4,5001159	4,2E-08	0,21449452	7E-12	c1.0001.0684 exported protein of unknown function 704928:705536 reverse
c1.0001.0685	0	-11,713218	2E-06	-6,1218063	2E-07	0,15377137	1E-10	c1.0001.0685 conserved exported protein of unknown function 705745:707334 reverse
c1.0001.0699	K00812	-6,5072164	3E-08	-3,4307961	1,4E-06	0,27554106	7E-10	c1.0001.0699 aspartate aminotransferase 725631:726836 forward
c1.0001.0710	K00384	-5,7973149	0,0151	-2,9454202	0,00966	0,3213222	2E-06	c1.0001.0710 thioredoxin reductase 736512:737606 reverse
c1.0001.0717	K01955	-8,9492151	0,0195	-4,5473772	0,00463	0,21020564	5E-12	c1.0001.0717 carbamoyl phosphate synthase large subunit 742713:745964 reverse
c1.0001.0721	0	-8,4015025	5E-16	-4,4547128	0,00075	0,20729711	5E-05	c1.0001.0721 diguanylate cyclase 748324:748815 reverse
c1.0001.0722	K03466	-7,9309318	2E-06	-4,2023866	5,8E-09	0,23027416	9E-13	c1.0001.0722 cell division protein FtsK 748849:750921 reverse
c1.0001.0723	0	-6,3033495	2E-05	-3,4352272	0,00464	0,30220438	1E-04	c1.0001.0723 hypothetical protein 750918:751631 reverse
c1.0001.0730	0	#DIV/0!	0,0005	#DIV/0!	0,02826	0,01419185	0,002	c1.0001.0730 short-chain dehydrogenase 756772:757755 reverse
c1.0001.0732	0	-17,44815	3E-05	-9,5936122	0,04979	0,12847504	0,016	c1.0001.0732 hypothetical protein 758529:758774 reverse
c1.0001.0783	K05805	-4,8092765	2E-12	-2,6169862	0,01837	0,36869126	0,002	c1.0001.0783 conserved exported protein of unknown function 809813:810346 forward
c1.0001.0806	0	-4,3967718	5E-26	-2,4356846	0,01241	0,3958436	5E-05	c1.0001.0806 transcriptional regulator 836562:837251 reverse

c1.0001.0811	0	#DIV/0!	8E-10	#DIV/0!	0,03136	0,01903316	0,007	c1.0001.0811 Hypothetical Protein 843432:843602 forward
c1.0001.0812	K02687	-5,650737	2E-06	-2,9574352	0,0125	0,31324442	0,002	c1.0001.0812 50S ribosomal protein L11 methyltransferase 843802:844716 forward
c1.0001.0828	K03589	-4,1000059	1E-05	-2,1856836	0,00346	0,4456918	9E-05	c1.0001.0828 F492457
c1.0001.0829	K01921	-4,0184892	0,0147	-2,0612822	0,02842	0,46337794	0,001	c1.0001.0829 D-alanine--D-alanine ligase 865272:866222 reverse
c1.0001.0830	K00075	#DIV/0!	4E-05	#DIV/0!	0,00249	0,01135712	1E-04	c1.0001.0830 UDP-N-acetylenolpyruvoylglucosamine reductase 866219:867145 reverse
c1.0001.0831	K01924	-7,4156931	0,0058	-3,867801	0,00187	0,25688234	2E-07	c1.0001.0831 UDP-N-acetylmuramate--alanine ligase 867159:868574 reverse
c1.0001.0838	K01929	-5,3148482	7E-12	-2,8592255	0,00032	0,33870064	1E-06	c1.0001.0838 UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate--D-alanyl-D-alanine ligase 873623:875119 reverse
c1.0001.0848	0	-13,711434	0,0002	-7,0830565	0,00838	0,13945694	0,001	c1.0001.0848 Hypothetical Protein 885892:886119 reverse
c1.0001.0851	K00102	-6,9604032	1E-10	-3,7154234	8,3E-07	0,25929848	2E-10	c1.0001.0851 lactate dehydrogenase 890765:892168 forward
c1.0001.0852	K00020	-8,9208254	6E-08	-4,8052613	0,00325	0,20323596	2E-04	c1.0001.0852 putative 3-hydroxyisobutyrate dehydrogenase 892458:893339 reverse
c1.0001.0853	0	-4,0834012	4E-10	-2,256785	0,03875	0,41843555	2E-04	c1.0001.0853 hypothetical protein 893703:893933 forward
c1.0001.0855	0	-4,081393	1E-04	-2,1477952	0,00901	0,44427378	4E-04	c1.0001.0855 histidine kinase 894313:896448 reverse
c1.0001.0864	K01878	-28,683712	6E-11	-14,906039	4,2E-08	0,06179761	9E-11	c1.0001.0864 glycyl-tRNA synthetase subunit alpha 901986:902873 forward
c1.0001.0866	K01649	-6,8008492	0,0004	-3,5407594	7,5E-08	0,26971665	2E-12	c1.0001.0866 2-isopropylmalate synthase 905172:906872 reverse
c1.0001.0892	K03786	-6,5847943	7E-28	-3,5172607	3,9E-05	0,26596082	2E-07	c1.0001.0892 3-dehydroquinate dehydratase 932500:932949 forward
c1.0001.0893	K02160	-5,3292091	6E-06	-2,7774727	0,00716	0,33482863	8E-04	c1.0001.0893 acetyl-CoA carboxylase 932973:933422 forward
c1.0001.0894	K01961	-10,01303	5E-09	-5,2934862	4,5E-09	0,17991161	4E-14	c1.0001.0894 acetyl-CoA carboxylase biotin carboxylase subunit 933443:934786 forward
c1.0001.0918	K00127	-5,5862258	1E-05	-2,9673673	0,0236	0,31048655	0,008	c1.0001.0918 formate dehydrogenase 955466:956512 forward
c1.0001.0924	K02109	-14,870725	0,0013	-7,7138684	2,1E-06	0,12753163	1E-09	c1.0001.0924 ATP synthase subunit B 961937:962416 reverse
c1.0001.0925	K02109	-39,982905	4E-24	-21,654122	2,3E-10	0,04709005	1E-14	c1.0001.0925 ATPase 962413:963063 reverse
c1.0001.0926	K02110	-18,318378	1E-187	-9,9634954	8,5E-11	0,09530896	1E-20	c1.0001.0926 ATP synthase subunit C 963151:963375 reverse
c1.0001.0927	K02108	-13,111172	7E-06	-6,8694847	8,4E-18	0,13896527	7E-26	c1.0001.0927 F0F1 ATP synthase subunit A 963449:964192 reverse
c1.0001.0928	K02116	-8,0538126	5E-29	-4,43335	2,9E-05	0,22638231	1E-07	c1.0001.0928 phosphoribosylaminoimidazolecarboxamide formyltransferase 964205:964567 reverse
c1.0001.0938	0	-11,212273	1E-05	-5,9108292	2,2E-06	0,1659505	4E-09	c1.0001.0938 chemotaxis protein 973575:974123 forward
c1.0001.0941	K19802	-4,3077599	3E-13	-2,3004782	0,00028	0,41423126	3E-06	c1.0001.0941 mandelate racemase 977027:978025 forward
c1.0001.0952	K01638	-6,5325418	3E-05	-3,4636371	5,9E-08	0,28003031	3E-10	c1.0001.0952 malate synthase 991539:993710 reverse
c1.0001.1019	K20035	-4,4963897	3E-08	-2,3857542	4,8E-05	0,39986039	1E-08	c1.0001.1019 acyl-CoA dehydrogenase 1059130:1060929 reverse

c1.0001.1022	0	-4,2852328	0,0005	-2,2622314	0,01142	0,43264477	7E-04	c1.0001.1022 2-hydroxyacid dehydrogenase 1062997:1064448 reverse
c1.0001.1027	K00982	-5,0686708	3E-08	-2,7264382	0,00014	0,35791501	7E-07	c1.0001.1027 glutamate-ammonia-ligase adenylyltransferase 1067912:1070929 reverse
c1.0001.1030	K01866	-4,9968147	2E-13	-2,7256793	0,00082	0,34868517	2E-05	c1.0001.1030 tyrosyl-tRNA synthetase 1074815:1076068 reverse
c1.0001.1037	K04755	-14,285518	2E-05	-8,131407	0,01717	0,12391849	9E-05	c1.0001.1037 2Fe-2S ferredoxin 1081658:1081987 forward
c1.0001.1038	K00566	-6,2240354	7E-10	-3,2560832	0,00064	0,27763009	5E-05	c1.0001.1038 thiouridylase 1082127:1083260 forward
c1.0001.1074	0	-7,2941732	5E-10	-4,0215256	0,00549	0,26235604	2E-04	c1.0001.1074 hypothetical protein 1123257:1123472 reverse
c1.0001.1079	0	-4,8022551	8E-54	-2,6037948	0,00195	0,36725843	5E-05	c1.0001.1079 hypothetical protein 1126902:1127078 forward
c1.0001.1081	K01589	-5,8420394	1E-05	-3,2018103	0,02833	0,31280903	0,005	c1.0001.1081 phosphoribosylaminoimidazole carboxilase 1129113:1130195 reverse
c1.0001.1083	0	-11,09942	1E-19	-6,3021841	0,00015	0,17361439	2E-06	c1.0001.1083 hypothetical protein ABAZ39
c1.0001.1093	0	-16,032285	0,0001	-8,4267371	1,8E-06	0,11879433	7E-09	c1.0001.1093 NAD(P)H quinone oxidoreductase 1139757:1140764 reverse
c1.0001.1110	K00773	-6,7632424	0,0003	-3,4832739	0,03046	0,25520321	0,011	c1.0001.1110 queuine tRNA-ribosyltransferase 1159051:1160253 reverse
c1.0001.1119	K01908	-5,2396896	6E-05	-2,7478579	6,4E-06	0,34756074	1E-08	c1.0001.1119 acyl-CoA synthetase 1170711:1172633 forward
c1.0001.1120	0	-4,9717376	3E-26	-2,6749341	0,00058	0,3572836	4E-06	c1.0001.1120 hypothetical protein ABAZ39
c1.0001.1131	K13503	-5,1796064	0,0006	-2,7379089	0,00025	0,35339523	7E-07	c1.0001.1131 RPE
c1.0001.1132	0	-15,82021	1E-10	-8,3485666	0,00261	0,11783579	3E-04	c1.0001.1132 hypothetical protein APCd
c1.0001.1133	K00324	-5,0710831	0,001	-2,6568629	9,2E-05	0,36263566	2E-07	c1.0001.1133 NAD(P) transhydrogenase subunit alpha 1188910:1190067 forward
c1.0001.1135	K00325	-4,1719486	0,0006	-2,1759246	0,00025	0,43788842	2E-06	c1.0001.1135 NAD synthetase 1190520:1191914 forward
c1.0001.1136	0	-7,4570375	2E-29	-4,1356776	0,00011	0,24119385	2E-06	c1.0001.1136 hypothetical protein ABAZ39
c1.0001.1144	0	-5,787207	0,0021	-2,9921862	0,01952	0,31619147	0,003	c1.0001.1144 hypothetical protein ABAZ39
c1.0001.1147	K01989	-13,607903	3E-05	-7,2214337	6,3E-06	0,14012818	2E-08	c1.0001.1147 ABC transporter substrate-binding protein 1200162:1201130 forward
c1.0001.1156	K01079	-26,721671	2E-05	-14,758966	0,00574	0,08390184	4E-04	c1.0001.1156 phosphoserine phosphatase 1210629:1211525 reverse
c1.0001.1160	K01652	-14,789035	1E-06	-7,7892545	5,7E-19	0,12320827	4E-27	c1.0001.1160 acetolactate synthase 1214788:1216551 forward
c1.0001.1161	K01653	-9,602963	0,0001	-4,8797301	2,8E-07	0,18540802	3E-10	c1.0001.1161 acetolactate synthase 1216602:1217108 forward
c1.0001.1162	K00053	-12,130281	3E-06	-6,3869716	8,2E-15	0,15072271	1E-20	c1.0001.1162 ketol-acid reductoisomerase 1217272:1218291 forward
c1.0001.1163	0	-5,8359711	0,0202	-2,961603	0,02072	0,32177165	7E-04	c1.0001.1163 PhoU family transcriptional regulator 1218485:1219189 reverse

c1.0001.1167	K03570	-5,4401788	1E-20	-2,9222189	0,00021	0,31988404	3E-06	c1.0001.1167 rod shape-determining protein MreC 1221511:1222380 forward
c1.0001.1186	K01998	-4,4653891	0,0001	-2,3435411	0,03872	0,40065377	0,007	c1.0001.1186 ABC transporter ATP-binding protein 1242276:1243991 forward
c1.0001.1222	K02952	-13,286804	0,0003	-6,879999	1,3E-09	0,13849504	4E-15	c1.0001.1222 30S ribosomal protein S13 1278262:1278630 forward
c1.0001.1223	K02948	-41,578142	0,0032	-21,691147	6E-08	0,04667719	5E-26	c1.0001.1223 MULTISPECIES: 30S ribosomal protein S11 1278646:1279044 forward
c1.0001.1224	K03040	-25,439171	0,0001	-13,099879	1,6E-15	0,07138275	2E-27	c1.0001.1224 DNA-directed RNA polymerase subunit alpha 1279176:1280195 forward
c1.0001.1225	K02879	-10,010583	6E-10	-5,3122343	1,8E-10	0,18084869	3E-15	c1.0001.1225 50S ribosomal protein L17 1280413:1280835 forward
c1.0001.1234	0	-7,942579	0,0056	-4,1060791	0,00206	0,23955451	2E-04	c1.0001.1234 benzoate transporter 1287147:1288358 forward
c1.0001.1240	K06140	-5,5674473	8E-48	-3,0988561	0,00892	0,31051614	1E-06	c1.0001.1240 nucleoside diphosphate kinase regulator 1292261:1292668 reverse
c1.0001.1271	0	-6,8239596	0,0269	-3,3923753	0,03231	0,27018309	5E-04	c1.0001.1271 histone acetyltransferase HPA2 1320728:1321144 forward
c1.0001.1273	0	-5,1149229	4E-08	-2,7113538	0,00029	0,35321857	8E-07	c1.0001.1273 dehydratase 1321991:1322446 forward
c1.0001.1284	K00330	-7,3390957	1E-06	-3,7992547	9,6E-07	0,24092705	4E-09	c1.0001.1284 NADH:ubiquinone oxidoreductase subunit A 1337649:1338017 forward
c1.0001.1285	K00331	-5,9196142	2E-14	-3,1768523	1,8E-05	0,3043778	2E-08	c1.0001.1285 NADH dehydrogenase 1338008:1338550 forward
c1.0001.1286	K00332	-5,858979	1E-16	-3,1212809	4,5E-05	0,30215565	1E-07	c1.0001.1286 NADH dehydrogenase subunit C 1338575:1339204 forward
c1.0001.1287	K00333	-5,9659705	0,0031	-3,0403129	0,0006	0,30955987	6E-08	c1.0001.1287 NADH dehydrogenase subunit D 1339201:1340421 forward
c1.0001.1288	K00334	-11,137542	8E-10	-5,7971079	3,7E-09	0,15725985	5E-13	c1.0001.1288 NADH dehydrogenase 1340418:1341035 forward
c1.0001.1289	K00335	-8,3956579	2E-15	-4,4494306	2,6E-09	0,21077242	8E-13	c1.0001.1289 NADH dehydrogenase 1341046:1342338 forward
c1.0001.1290	K00336	-11,587802	6E-07	-6,1191822	1,5E-15	0,15698889	1E-22	c1.0001.1290 NADH dehydrogenase subunit G 1342354:1344438 forward
c1.0001.1291	K00337	-11,055404	4E-19	-5,9389484	1,2E-13	0,16261627	2E-19	c1.0001.1291 NADH:ubiquinone oxidoreductase subunit H 1344431:1345441 forward
c1.0001.1292	K00338	-9,495431	1E-12	-5,0349184	3,3E-09	0,18857772	5E-14	c1.0001.1292 NADH dehydrogenase subunit I 1345536:1346024 forward
c1.0001.1293	K00339	-17,826059	1E-06	-9,4023415	2,5E-13	0,10354805	3E-19	c1.0001.1293 NADH:ubiquinone oxidoreductase subunit J 1346061:1346675 forward
c1.0001.1294	K00340	-19,072203	3E-11	-10,351587	1,5E-10	0,09977009	5E-16	c1.0001.1294 NADH-quinone oxidoreductase 1346678:1346986 forward

c1.0001.1295	K00341	-10,465796	1E-05	-5,5009888	1,2E-14	0,17405755	1E-22	c1.0001.1295 NADH:ubiquinone oxidoreductase subunit L 1346996:1349083 forward
c1.0001.1296	K00342	-13,625182	0,0001	-7,0813785	5E-14	0,13434262	3E-25	c1.0001.1296 NADH:ubiquinone oxidoreductase subunit M 1349108:1350634 forward
c1.0001.1297	K00343	-7,6255576	2E-12	-4,0793098	1,6E-10	0,23564479	8E-16	c1.0001.1297 NADH:ubiquinone oxidoreductase subunit N 1350645:1352093 forward
c1.0001.1298	K03524	-10,743117	4E-05	-5,7889417	0,02471	0,16850201	0,005	c1.0001.1298 biotin 1352090:1352884 forward
c1.0001.1306	K01881	-8,7592217	0,0016	-4,4980501	5,2E-06	0,2097098	5E-12	c1.0001.1306 proline--tRNA ligase 1358536:1359849 forward
c1.0001.1314	K02967	-4,2315717	1E-08	-2,2441547	0,00015	0,42468223	5E-07	c1.0001.1314 30S ribosomal protein S2 1368581:1369375 forward
c1.0001.1315	K02357	-6,8226401	5E-07	-3,5935203	4,8E-09	0,26514731	1E-13	c1.0001.1315 elongation factor Ts 1369491:1370420 forward
c1.0001.1316	K09903	-22,783759	3E-05	-13,024566	0,00121	0,07455073	2E-20	c1.0001.1316 uridylate kinase 1370475:1371230 forward
c1.0001.1317	K02838	-5,5045871	5E-08	-2,8722901	0,00243	0,31831606	2E-04	c1.0001.1317 ribosome recycling factor 1371282:1371803 forward
c1.0001.1318	K00806	-5,2224898	0,001	-2,6868276	0,0264	0,3461929	0,006	c1.0001.1318 UDP pyrophosphate synthase 1371808:1372545 forward
c1.0001.1322	K07277	-5,2067423	1E-14	-2,8185094	2,9E-06	0,34205509	3E-08	c1.0001.1322 membrane protein 1375754:1378120 forward
c1.0001.1323	0	-4,1485692	9E-17	-2,285479	0,00715	0,42793926	3E-04	c1.0001.1323 membrane protein 1378122:1378700 forward
c1.0001.1324	K02372	-5,3404782	6E-12	-2,8327932	9,9E-05	0,33365738	2E-07	c1.0001.1324 3-hydroxyacyl-ACP dehydratase 1378915:1379385 forward
c1.0001.1325	K00677	-4,1338292	0,0009	-2,1343029	0,00402	0,43815742	3E-04	c1.0001.1325 UDP-N-acetylglucosamine acyltransferase 1379382:1380170 forward
c1.0001.1331	K00799	-8,0697282	1E-08	-4,2828732	2E-09	0,22474841	1E-13	c1.0001.1331 glutathione S-transferase 1384307:1384915 forward
c1.0001.1332	K01647	-3,7706293	3E-08	-2,0006159	0,00205	0,47617148	1E-06	c1.0001.1332 type II citrate synthase 1385119:1386432 reverse
c1.0001.1333	K01885	-7,5520652	0,0004	-3,9375917	1,4E-06	0,24459168	1E-09	c1.0001.1333 glutamyl-tRNA synthetase 1386527:1387939 reverse
c1.0001.1336	K03750	-5,6589957	3E-05	-3,0752222	0,01769	0,29886805	0,004	c1.0001.1336 molybdenum cofactor biosynthesis protein MoaA 1391177:1392394 reverse
c1.0001.1338	K03829	-4,3424565	6E-15	-2,3516829	0,01805	0,39687697	0,003	c1.0001.1338 AAT
c1.0001.1339	K01609	-5,4956593	1E-05	-2,8987643	0,00938	0,33207751	8E-04	c1.0001.1339 RPC
c1.0001.1340	K00766	-4,1584934	1E-07	-2,2965843	0,0299	0,4219758	0,006	c1.0001.1340 RPD
c1.0001.1341	K01658	-5,1504012	7E-22	-2,7595614	0,00187	0,33127979	4E-05	c1.0001.1341 anthranilate synthase 1395237:1395827 reverse
c1.0001.1345	0	-18,191923	0,0002	-10,268337	0,03633	0,11930268	0,011	c1.0001.1345 hypothetical protein ABAZ39
c1.0001.1347	K03770	-21,164666	2E-05	-11,125567	1,5E-10	0,08834139	2E-15	c1.0001.1347 peptidylprolyl isomerase 1400355:1402241 reverse
c1.0001.1348	K01803	-10,393895	2E-08	-5,686704	0,00346	0,18120966	2E-04	c1.0001.1348 triosephosphate isomerase 1402470:1403225 forward
c1.0001.1349	K03075	-6,204135	2E-06	-3,292824	0,00043	0,29680398	5E-06	c1.0001.1349 preprotein translocase subunit SecG 1403354:1403686 forward
c1.0001.1350	K01937	-9,2295074	2E-07	-4,8159361	1,6E-10	0,19347552	6E-14	c1.0001.1350 CTP synthetase 1403847:1405475 forward

c1.0001.1360	0	-8,9708544	4E-16	-4,8261755	1,8E-05	0,19186033	3E-07	c1.0001.1360 hypothetical protein ABAZ39
c1.0001.1363	0	-8,4544275	3E-05	-4,4596131	1,1E-08	0,21830053	4E-12	c1.0001.1363 membrane protein 1417459:1418250 forward
c1.0001.1364	0	-5,9377657	0,0064	-3,0672349	0,00198	0,31391476	7E-08	c1.0001.1364 membrane protein 1418358:1418582 forward
c1.0001.1365	0	-5,5000778	4E-33	-2,9485514	2,9E-07	0,3223608	3E-11	c1.0001.1365 peroxiredoxin 1418640:1419170 reverse
c1.0001.1371	0	-4,01106	3E-34	-2,1582937	0,00028	0,43912816	6E-06	c1.0001.1371 hypothetical protein ABAZ39
c1.0001.1372	0	-18,09476	2E-07	-10,115248	0,0024	0,11496021	1E-04	c1.0001.1372 esterase 1427455:1428234 reverse
c1.0001.1374	K01011	-10,393418	0,002	-5,3832741	0,0013	0,18303878	1E-04	c1.0001.1374 sulfurtransferase 1429961:1430803 reverse
c1.0001.1377	K01751	-3,9496606	5E-11	-2,0909078	0,00614	0,43926505	0,001	c1.0001.1377 diaminopropionate ammonia-lyase 1431878:1433110 forward
c1.0001.1378	0	-4,8280916	3E-10	-2,6121946	0,0049	0,36924617	2E-04	c1.0001.1378 aspartate aminotransferase 1433213:1434394 forward
c1.0001.1402	K01740	-7,9601899	2E-22	-4,3246716	6,1E-07	0,22594851	1E-10	c1.0001.1402 O-acetylhomoserine aminocarboxypropyltransferase 1461361:1462650 reverse
c1.0001.1403	K06929	-8,3182131	0,0009	-4,2221219	0,00216	0,21504312	4E-04	c1.0001.1403 CoA-binding protein 1462669:1463187 reverse
c1.0001.1404	0	-15,196178	7E-12	-8,4937997	6,2E-05	0,12114241	1E-06	c1.0001.1404 enoyl-CoA hydratase 1463184:1463990 reverse
c1.0001.1407	K02871	-66,390988	0,0064	-33,878676	4,8E-08	0,02843899	2E-37	c1.0001.1407 50S ribosomal protein L13 1464917:1465381 forward
c1.0001.1408	K02996	-24,951104	6E-07	-12,942358	5,7E-14	0,07182812	3E-20	c1.0001.1408 30S ribosomal protein S9 1465386:1465868 forward
c1.0001.1412	K03821	-6,2555167	0,0004	-3,2634091	6,2E-07	0,29227757	2E-13	c1.0001.1412 poly(3-hydroxyalkanoate) synthetase 1468589:1470388 reverse
c1.0001.1413	0	-4,3687612	8E-05	-2,2786057	0,00736	0,41379677	4E-04	c1.0001.1413 hypothetical protein 1470622:1471017 forward
c1.0001.1417	0	-7,20537	0,0023	-3,7215874	4,8E-05	0,25684419	2E-10	c1.0001.1417 Hypothetical Protein 1473212:1473871 reverse
c1.0001.1419	K00003	-4,553435	0,0001	-2,395034	0,00077	0,40226597	8E-06	c1.0001.1419 homoserine dehydrogenase 1475460:1476767 forward
c1.0001.1425	K01154	-4,6712081	1E-06	-2,4442534	1,2E-05	0,3868362	8E-08	c1.0001.1425 Hypothetical Protein 1485574:1487301 forward
c1.0001.1426	0	-5,1995554	5E-08	-2,7367022	6,2E-06	0,34510617	3E-08	c1.0001.1426 Hypothetical Protein 1487427:1488386 forward
c1.0001.1433	0	-5,4235363	2E-07	-2,8714643	0,00011	0,33442809	2E-07	c1.0001.1433 Hypothetical Protein 1493692:1494180 reverse
c1.0001.1434	0	-5,1332884	1E-07	-2,6890764	1,9E-06	0,34971955	3E-08	c1.0001.1434 hypothetical protein 1494335:1496185 reverse
c1.0001.1435	0	-7,2446231	0,0333	-3,6284129	0,02322	0,25982966	3E-07	c1.0001.1435 Hypothetical Protein 1496260:1496757 reverse
c1.0001.1436	0	-7,6253247	5E-07	-3,9920641	8,5E-11	0,23715578	8E-17	c1.0001.1436 Hypothetical Protein 1496741:1497844 reverse
c1.0001.1488	K03188	-5,1586608	5E-06	-2,7181606	0,01816	0,35028111	0,003	c1.0001.1488 urease accessory protein UreF 1553613:1554359 forward
c1.0001.1507	0	-5,2786986	1E-08	-2,773685	0,00198	0,33559182	7E-05	c1.0001.1507 transcriptional regulator 1571554:1572078 forward
c1.0001.1511	0	-4,1590496	9E-06	-2,1936298	0,00046	0,43575371	5E-06	c1.0001.1511 signal peptide protein 1574830:1575639 reverse
c1.0001.1513	K00721	-8,284518	0,0027	-4,5019677	0,02174	0,24728916	0,002	c1.0001.1513 dolichol-phosphate mannosyltransferase 1577860:1578588 forward

c1.0001.1515	0	-5,4730584	3E-07	-2,9608803	0,00323	0,32695832	2E-04	c1.0001.1515 glycosyl transferase 1578940:1580613 forward
c1.0001.1533	0	-11,815249	2E-05	-6,5185908	0,00905	0,17461374	5E-04	c1.0001.1533 oxidoreductase 1595538:1596248 reverse
c1.0001.1534	K00764	-4,0902045	0,0038	-2,1384551	0,00763	0,45265602	1E-04	c1.0001.1534 amidophosphoribosyltransferase 1596339:1597787 reverse
c1.0001.1537	0	-4,3527147	0,0028	-2,2831487	0,00598	0,42908515	3E-04	c1.0001.1537 methionine aminopeptidase 1600199:1601026 forward
c1.0001.1539	0	-5,1344723	6E-25	-2,835085	0,0015	0,34730497	5E-06	c1.0001.1539 hypothetical protein ABAZ39
c1.0001.1540	K02065	-4,4725844	9E-08	-2,4066196	0,04882	0,40843365	0,011	c1.0001.1540 organic solvent ABC transporter ATP-binding protein 1601939:1602730 reverse
c1.0001.1541	K02066	-4,4020798	0,0002	-2,3536643	0,02669	0,42588623	0,003	c1.0001.1541 ABC transporter permease 1602739:1603515 reverse
c1.0001.1542	K01775	-4,6326101	1E-12	-2,4882422	0,00209	0,37212859	2E-04	c1.0001.1542 alanine racemase 1603515:1604651 reverse
c1.0001.1545	K02939	-11,083111	2E-05	-5,8421176	2,2E-09	0,16740802	3E-14	c1.0001.1545 50S ribosomal protein L9 1607849:1608415 reverse
c1.0001.1546	0	-42,538966	1E-05	-21,964471	1,2E-24	0,04240711	6E-31	c1.0001.1546 membrane protein 1608415:1609371 reverse
c1.0001.1547	K02963	-17,260328	9E-05	-8,8733485	1,4E-09	0,10491718	4E-13	c1.0001.1547 30S ribosomal protein S18 1609406:1609678 reverse
c1.0001.1548	K02990	-35,716065	6E-06	-19,123464	2,3E-23	0,05357303	3E-29	c1.0001.1548 30S ribosomal protein S6 1609679:1610116 reverse
c1.0001.1550	K00645	-8,0565078	0,005	-4,1304448	0,00054	0,23180485	2E-05	c1.0001.1550 ACP S-malonyltransferase 1610510:1611466 forward
c1.0001.1552	K02078	-9,0005516	6E-09	-4,8377183	1,1E-09	0,20410309	7E-13	c1.0001.1552 acyl carrier protein 1612382:1612621 forward
c1.0001.1553	K09458	-4,2239918	1E-19	-2,283627	0,00072	0,41096446	3E-05	c1.0001.1553 3-oxoacyl-ACP synthase 1612709:1613971 forward
c1.0001.1557	K01714	-7,1342874	1E-11	-3,7877962	1,8E-08	0,25123172	3E-12	c1.0001.1557 dihydrodipicolinate synthase 1617583:1618458 forward
c1.0001.1558	K03664	-7,7378637	0,0014	-4,0160385	0,00083	0,24221486	4E-05	c1.0001.1558 single-stranded DNA-binding protein 1618495:1618971 forward
c1.0001.1562	K01698	-7,7680358	6E-28	-4,1702619	5,6E-08	0,22751516	7E-12	c1.0001.1562 delta-aminolevulinic acid dehydratase 1621604:1622617 reverse
c1.0001.1565	K01692	-7,2909583	2E-06	-3,8426127	0,00735	0,24623121	7E-04	c1.0001.1565 enoyl-CoA hydratase 1624690:1625781 forward
c1.0001.1574	0	-9,1295462	2E-07	-5,1664473	0,0278	0,19416425	0,007	c1.0001.1574 hypothetical protein 1633956:1634522 forward
c1.0001.1600	K01496	-5,2451618	3E-09	-2,9046672	0,03363	0,3497359	0,006	c1.0001.1600 phosphoribosyl-AMP cyclohydrolase 1658142:1658576 reverse
c1.0001.1601	K01760	-4,601209	4E-08	-2,4741257	0,00274	0,39625498	5E-05	c1.0001.1601 cystathionine beta-lyase 1658619:1659797 reverse
c1.0001.1612	K01738	-8,5463911	8E-15	-4,5414755	1,1E-05	0,20647136	3E-08	c1.0001.1612 cysteine synthase 1669242:1670246 reverse
c1.0001.1618	K00573	-4,1477009	0,0003	-2,1724421	0,00089	0,44051823	2E-05	c1.0001.1618 protein-L-isoaspartate O-methyltransferase 1673874:1674530 forward
c1.0001.1621	K09991	-3,9633647	0,0002	-2,0717528	0,02703	0,45784835	0,003	c1.0001.1621 hypothetical protein ABAZ39
c1.0001.1622	K01873	-5,8048197	2E-13	-3,1016428	9E-07	0,30698269	6E-10	c1.0001.1622 valyl-tRNA synthetase 1677235:1679886 forward
c1.0001.1626	K07258	-5,9739669	0,0001	-3,1095518	1,9E-06	0,30487805	7E-09	c1.0001.1626 D-alanyl-D-alanine carboxypeptidase 1684038:1685204 forward

c1.0001.1628	K02341	-6,9886446	6E-05	-3,9090478	0,02309	0,28022439	0,002	c1.0001.1628 DNA polymerase III subunit delta' 1685909:1687027 forward
c1.0001.1630	K03424	-5,0970777	3E-06	-2,6853011	0,01232	0,35342047	0,001	c1.0001.1630 LuxR family transcriptional regulator 1688657:1689454 forward
c1.0001.1637	K11074	-6,0536531	3E-12	-3,3311627	0,00147	0,30086135	5E-05	c1.0001.1637 putrescine/spermidine ABC transporter permease 1695416:1696249 reverse
c1.0001.1652	K03823	-4,1505927	2E-07	-2,1811296	0,02831	0,42635915	0,005	c1.0001.1652 GCN5 family N-acetyltransferase 1710649:1711173 reverse
c1.0001.1658	0	-5,1448933	9E-06	-2,8003926	0,02387	0,3392689	0,005	c1.0001.1658 hypothetical protein ABAZ39
c1.0001.1659	K05834	-5,3398507	3E-14	-2,8713864	0,00015	0,33736125	3E-07	c1.0001.1659 amino acid transporter 1717155:1717790 reverse
c1.0001.1664	K03753	-5,9940059	0,0028	-3,033008	0,00787	0,29850718	0,002	c1.0001.1664 molybdopterin-guanine dinucleotide biosynthesis protein MobB 1720659:1721186 reverse
c1.0001.1666	K03703	-4,1996836	1E-05	-2,2681378	0,0089	0,43631005	3E-04	c1.0001.1666 excinuclease ABC subunit C 1722012:1723994 reverse
c1.0001.1674	K01676	-10,363937	2E-08	-5,4257385	3E-13	0,17254566	2E-17	c1.0001.1674 fumarate hydratase 1736612:1738246 forward
c1.0001.1675	K03977	-5,2909052	1E-16	-2,8645294	0,00017	0,33625224	5E-07	c1.0001.1675 GTP-binding protein Der 1738348:1739751 reverse
c1.0001.1676	0	-13,903942	0,0003	-7,4510373	9E-06	0,14213274	4E-08	c1.0001.1676 pyrrolo-quinoline quinone 1739788:1741170 reverse
c1.0001.1686	K03549	-4,3446745	0,001	-2,287521	0,01364	0,42695059	0,001	c1.0001.1686 potassium transporter Kup 1748838:1750790 forward
c1.0001.1694	K00808	-4,2834317	1E-05	-2,2482993	0,00012	0,42244841	6E-07	c1.0001.1694 homospermidine synthase 1761285:1762703 reverse
c1.0001.1706	K02049	-7,6587698	0,0072	-3,9037849	0,007	0,24422158	0,001	c1.0001.1706 ABC transporter 1772964:1773716 reverse
c1.0001.1707	K02050	-11,474795	1E-05	-6,2536767	0,01608	0,15582381	0,003	c1.0001.1707 nitrate ABC transporter permease 1773704:1774468 reverse
c1.0001.1708	K02051	-11,142986	2E-08	-6,5365406	0,02441	0,16724007	4E-05	c1.0001.1708 nitrate ABC transporter substrate-binding protein 1774474:1775520 reverse
c1.0001.1709	K00955	-15,373137	4E-16	-8,1643319	4,1E-18	0,11585768	3E-25	c1.0001.1709 NodQ bifunctional enzyme 1775717:1777591 reverse
c1.0001.1710	K00957	-30,481953	0,0012	-15,68809	1,2E-09	0,06103504	2E-25	c1.0001.1710 sulfate adenylyltransferase 1777610:1778407 reverse
c1.0001.1711	K00390	-101,39222	6E-05	-55,350385	1,4E-09	0,02257197	7E-13	c1.0001.1711 phosphoadenosine phosphosulfate reductase 1778404:1779138 reverse
c1.0001.1714	K02303	-6,0307152	0,0247	-2,9871974	0,04697	0,30205187	0,001	c1.0001.1714 uroporphyrin-III methyltransferase 1780030:1780725 forward
c1.0001.1716	K00381	-4,771853	9E-11	-2,5516591	5,1E-05	0,37608721	2E-07	c1.0001.1716 sulfite reductase 1781133:1782821 forward
c1.0001.1717	0	-9,6302409	0,0006	-4,771703	0,0018	0,17570597	2E-04	c1.0001.1717 hypothetical protein 1782808:1783305 forward
c1.0001.1723	K00648	-7,9376125	0,0116	-3,9391035	0,00766	0,22723415	2E-05	c1.0001.1723 3-oxoacyl-ACP synthase 1788788:1789759 reverse
c1.0001.1724	K03621	-7,2998812	2E-25	-3,9171347	7,1E-07	0,24176408	2E-10	c1.0001.1724 phosphate acyltransferase 1789764:1790816 reverse
c1.0001.1725	K02911	-14,969726	3E-09	-8,4587165	0,00074	0,11679158	2E-13	c1.0001.1725 50S ribosomal protein L32 1790866:1791051 reverse

c1.0001.1732	K03625	-5,5891308	0,0035	-2,9268202	0,02179	0,34195806	0,003	c1.0001.1732 transcription antitermination protein NusB 1797021:1797593 reverse
c1.0001.1733	K00794	-6,1929845	0,0109	-3,1255627	0,00605	0,29875624	7E-07	c1.0001.1733 6,7-dimethyl-8-ribityllumazine synthase 1797627:1798091 reverse
c1.0001.1734	K14652	-4,0609566	9E-11	-2,1460199	0,00307	0,43320716	2E-04	c1.0001.1734 3,4-dihydroxy-2-butanone 4-phosphate synthase 1798113:1799237 reverse
c1.0001.1735	K00793	-8,4053332	1E-22	-4,5996934	0,00056	0,20356047	9E-07	c1.0001.1735 riboflavin synthase subunit alpha 1799234:1799821 reverse
c1.0001.1758	K06207	-10,686762	9E-05	-5,525217	6,8E-11	0,16944536	2E-14	c1.0001.1758 GTP-binding protein TypA 1821534:1823360 forward
c1.0001.1759	0	-5,5985673	0,02	-2,7956316	0,03631	0,3270862	5E-04	c1.0001.1759 peroxiredoxin 1823477:1823950 forward
c1.0001.1775	0	-7,4794404	4E-38	-3,9978377	1,5E-06	0,22921724	3E-09	c1.0001.1775 hypothetical protein 1835011:1835601 reverse
c1.0001.1780	K02231	-11,074927	0,0001	-5,6145178	0,00093	0,15913057	8E-05	c1.0001.1780 adenosylcobinamide kinase 1841870:1842394 reverse
c1.0001.1782	0	-6,8418185	1E-10	-3,7140966	0,00714	0,27337957	4E-04	c1.0001.1782 hypothetical protein ABAZ39
c1.0001.1790	K03723	-4,2937657	7E-06	-2,2949524	0,0034	0,42313881	5E-05	c1.0001.1790 transcription-repair coupling factor 1849593:1853219 reverse
c1.0001.1791	K09159	-5,8531588	3E-08	-3,0346822	0,01468	0,29305675	0,004	c1.0001.1791 hypothetical protein 1853230:1853592 reverse
c1.0001.1798	K00383	-5,7025672	0,0002	-2,976042	7E-05	0,32063669	2E-07	c1.0001.1798 glutathione reductase 1859355:1860707 forward
c1.0001.1799	K07217	-5,8581279	9E-05	-3,072771	0,00157	0,31328371	4E-05	c1.0001.1799 Mn-containing catalase 1860878:1861774 forward
c1.0001.1800	K08218	-6,3226182	1E-05	-3,3865279	0,01039	0,26194904	0,002	c1.0001.1800 MFS transporter permease 1861917:1863260 reverse
c1.0001.1801	K01626	-5,4583933	1E-15	-2,9215914	2,8E-06	0,32686498	3E-09	c1.0001.1801 phospho-2-dehydro-3-deoxyheptonate aldolase 1863476:1864861 forward
c1.0001.1804	K01918	-4,443573	2E-06	-2,3748784	0,02305	0,41293676	0,003	c1.0001.1804 pantoate--beta-alanine ligase 1867211:1868101 reverse
c1.0001.1807	0	-4,4128439	9E-05	-2,3225043	0,0237	0,41658552	0,003	c1.0001.1807 hypothetical protein 1870004:1870282 reverse
c1.0001.1811	0	-14,041586	7E-08	-7,3666632	0,00013	0,13002444	3E-06	c1.0001.1811 membrane protein 1874740:1875267 reverse
c1.0001.1821	0	-12,722504	7E-79	-6,8614092	1,1E-06	0,1371754	4E-09	c1.0001.1821 Hypothetical Protein 1882814:1882921 forward
c1.0001.1822	0	-13,631655	0,0059	-7,917185	0,0383	0,12284035	3E-10	c1.0001.1822 hypothetical protein 1882974:1883294 forward
c1.0001.1827	K11689	-6,1043127	0,0284	-3,0590327	0,03211	0,30532084	1E-04	c1.0001.1827 C4-dicarboxylate ABC transporter substrate-binding protein 1889128:1889826 forward
c1.0001.1839	K00297	-4,7143112	0,0022	-2,4716416	0,02636	0,3987848	0,004	c1.0001.1839 5,10-methylenetetrahydrofolate reductase 1904076:1904957 reverse
c1.0001.1843	0	-4,6622549	1E-05	-2,4603492	0,00129	0,38945078	1E-05	c1.0001.1843 hypothetical protein ABAZ39
c1.0001.1848	0	-4,2886388	2E-19	-2,3136886	0,00033	0,41579519	4E-06	c1.0001.1848 hypothetical protein ABAZ39
c1.0001.1885	0	-3,9022451	1E-17	-2,0650092	0,00107	0,45048451	3E-04	c1.0001.1885 Hypothetical Protein 1955091:1955471 forward
c1.0001.1927	0	#DIV/0!	8E-24	#DIV/0!	0,00014	0,00765963	9E-06	c1.0001.1927 hypothetical protein 1995880:1996029 reverse

c1.0001.1928	K01953	-6,0863794	5E-06	-3,2992278	0,00181	0,28348241	1E-04	c1.0001.1928 asparagine synthase 1996026:1997993 reverse
c1.0001.1929	0	-12,077085	4E-36	-6,5253692	2,6E-06	0,14419633	2E-08	c1.0001.1929 hypothetical protein 1998011:1998349 reverse
c1.0001.1936	K03640	-17,610988	0	-9,5965752	2,1E-16	0,0997295	9E-34	c1.0001.1936 outer membrane lipoprotein of the Pal-Tol system 2005227:2005703 reverse
c1.0001.1937	K03641	-17,76372	3E-11	-9,2715564	1,4E-09	0,0982544	1E-13	c1.0001.1937 translocation protein TolB 2005852:2007204 reverse
c1.0001.1938	0	-6,6889073	1E-06	-3,6943949	0,00642	0,25903888	0,002	c1.0001.1938 membrane protein 2007207:2008307 reverse
c1.0001.1939	K03560	-12,590086	0,0033	-6,3335608	0,00197	0,14617764	4E-04	c1.0001.1939 biopolymer transporter ExbD 2008393:2008875 reverse
c1.0001.1950	K00615	-13,11893	7E-14	-7,0813114	1,5E-10	0,13937372	4E-16	c1.0001.1950 transketolase 2016380:2018392 forward
c1.0001.1951	K00134	-17,357958	1E-51	-9,4058793	2E-22	0,10257268	5E-32	c1.0001.1951 glyceraldehyde-3-phosphate dehydrogenase 2018537:2019544 forward
c1.0001.1961	K02356	-31,492913	1E-230	-17,250847	3,5E-11	0,05530276	2E-33	c1.0001.1961 elongation factor P 2027351:2027917 forward
c1.0001.1970	0	-4,992899	7E-05	-2,5579041	0,046	0,34870931	0,018	c1.0001.1970 glutamate synthase 2037376:2037621 reverse
c1.0001.1974	K03469	#DIV/0!	0,0177	#DIV/0!	0,00209	0,00855369	2E-05	c1.0001.1974 ribonuclease H 2040381:2040875 forward
c1.0001.1982	K01733	-6,6918297	0,0003	-3,4978476	0,00014	0,27599457	1E-06	c1.0001.1982 threonine synthase 2048347:2049759 reverse
c1.0001.2009	K01939	-4,9332667	8E-05	-2,5588895	0,00024	0,36428151	4E-06	c1.0001.2009 adenylosuccinate synthetase 2077375:2078667 reverse
c1.0001.2013	K00831	-6,7358386	2E-18	-3,6548736	1,9E-05	0,26132575	5E-08	c1.0001.2013 phosphoserine aminotransferase 2083572:2084732 reverse
c1.0001.2020	0	-7,7501315	1E-07	-4,3660198	0,0211	0,23387245	0,004	c1.0001.2020 hypothetical protein 2091965:2092534 forward
c1.0001.2037	K02355	-14,279268	8E-06	-7,4839439	1,8E-16	0,12742944	4E-33	c1.0001.2037 elongation factor G 2112316:2114460 reverse
c1.0001.2038	0	-8,7406956	4E-10	-4,6320547	1,1E-09	0,20630978	4E-14	c1.0001.2038 endoribonuclease 2114690:2115037 reverse
c1.0001.2056	K01956	-6,3488365	9E-06	-3,3082366	4,9E-05	0,28218295	3E-07	c1.0001.2056 carbamoyl phosphate synthase small subunit 2135876:2137051 forward
c1.0001.2060	K00058	-5,0241828	8E-05	-2,6417539	4,6E-05	0,3636066	2E-07	c1.0001.2060 3-phosphoglycerate dehydrogenase 2141557:2142804 reverse
c1.0001.2063	0	-4,8667836	0,0003	-2,5627264	0,0085	0,38157917	5E-04	c1.0001.2063 GCN5 family acetyltransferase 2144320:2144910 reverse
c1.0001.2066	K01613	-5,9883586	0,0015	-3,0923288	0,00018	0,30702976	3E-06	c1.0001.2066 phosphatidylserine decarboxylase 2146410:2147108 reverse
c1.0001.2071	K00031	-6,7822847	0,0002	-3,5357002	8,8E-09	0,26951077	9E-15	c1.0001.2071 isocitrate dehydrogenase 2150142:2151359 forward
c1.0001.2095	K02401	-6,0764991	0,0032	-3,1600108	0,00056	0,30822972	1E-06	c1.0001.2095 flagellar biosynthesis protein 2175024:2176100 reverse
c1.0001.2108	0	-3,9246348	2E-06	-2,0910067	0,02552	0,46467434	0,002	c1.0001.2108 hypothetical protein ABAZ39
c1.0001.2134	K06147	-4,1322496	5E-06	-2,2272006	0,00841	0,44376207	3E-04	c1.0001.2134 ABC transporter 2210302:2212104 reverse
c1.0001.2135	K02909	-10,9717	6E-05	-5,6730304	7,7E-12	0,16513328	3E-15	c1.0001.2135 50S ribosomal protein L31 2212468:2212692 forward

c1.0001.2143	0	-3,9863605	0,0028	-2,0892423	0,00716	0,46560075	4E-04	c1.0001.2143 membrane protein 2220496:2221410 forward
c1.0001.2144	0	-4,9020209	4E-07	-2,5920938	9,8E-05	0,36874178	5E-07	c1.0001.2144 transcriptional regulator 2221502:2222278 forward
c1.0001.2146	K03076	-6,858311	7E-10	-3,6352361	1,8E-09	0,26211226	7E-16	c1.0001.2146 preprotein translocase subunit SecY 2223059:2224399 reverse
c1.0001.2147	K02876	-7,7205214	8E-40	-4,1596795	4,9E-07	0,22903293	2E-10	c1.0001.2147 50S ribosomal protein L15 2224518:2225003 reverse
c1.0001.2148	K02907	-18,745216	9E-33	-10,334576	3,9E-05	0,09804876	1E-06	c1.0001.2148 50S ribosomal protein L30 2225008:2225199 reverse
c1.0001.2149	K02988	-12,366979	3E-05	-6,5470574	1,7E-15	0,14906573	2E-19	c1.0001.2149 30S ribosomal protein S5 2225202:2225783 reverse
c1.0001.2150	K02881	-7,8845008	3E-08	-4,1374657	3E-09	0,22744905	9E-13	c1.0001.2150 50S ribosomal protein L18 2225796:2226158 reverse
c1.0001.2151	K02933	-7,3052462	2E-10	-4,0400321	0,00317	0,23426744	2E-11	c1.0001.2151 50S ribosomal protein L6 2226167:2226700 reverse
c1.0001.2152	K02994	-8,9593098	8E-47	-4,8254578	1,1E-10	0,19784492	5E-15	c1.0001.2152 30S ribosomal protein S8 2226714:2227112 reverse
c1.0001.2153	K02954	-16,2376	4E-13	-8,6168027	4,3E-21	0,11063613	1E-31	c1.0001.2153 30S ribosomal protein S14 2227126:2227431 reverse
c1.0001.2154	K02931	-12,834052	4E-06	-6,6676969	1E-14	0,14196367	7E-26	c1.0001.2154 50S ribosomal protein L5 2227451:2227993 reverse
c1.0001.2155	K02895	-12,795799	9E-11	-6,7284408	4,7E-18	0,14007045	9E-26	c1.0001.2155 50S ribosomal protein L24 2228005:2228322 reverse
c1.0001.2156	K02874	-6,8869617	5E-41	-3,7154954	1,4E-09	0,25759848	2E-14	c1.0001.2156 MULTISPECIES: 50S ribosomal protein L14 2228322:2228690 reverse
c1.0001.2157	K02961	-10,665372	1E-107	-5,747247	2E-13	0,16555659	9E-27	c1.0001.2157 30S ribosomal protein S17 2228708:2228962 reverse
c1.0001.2158	K02904	-22,415276	8E-162	-12,05457	2,6E-08	0,07514819	2E-19	c1.0001.2158 50S ribosomal protein L29 2228975:2229181 reverse
c1.0001.2159	K02878	-43,162121	2E-06	-23,017434	6,8E-20	0,04439818	1E-27	c1.0001.2159 50S ribosomal protein L16 2229178:2229606 reverse
c1.0001.2160	K02982	-19,6032	0,0027	-10,081691	2,3E-07	0,09475582	8E-31	c1.0001.2160 30S ribosomal protein S3 2229622:2230398 reverse
c1.0001.2161	K02890	-27,623992	3E-05	-14,26998	9,6E-25	0,06576389	4E-40	c1.0001.2161 50S ribosomal protein L22 2230400:2230795 reverse
c1.0001.2163	K02886	-64,78606	4E-05	-33,812511	1,2E-29	0,02838268	4E-63	c1.0001.2163 50S ribosomal protein L2 2231090:2231917 reverse
c1.0001.2164	K02892	-43,662337	0,0014	-22,504631	6E-11	0,04271405	3E-38	c1.0001.2164 50S ribosomal protein L23 2231934:2232251 reverse
c1.0001.2165	K02926	-57,741606	4E-13	-30,393294	1,9E-40	0,03087552	3E-59	c1.0001.2165 50S ribosomal protein L4 2232248:2232868 reverse
c1.0001.2166	K02906	-29,658029	2E-43	-15,8311	6,4E-29	0,05946235	6E-39	c1.0001.2166 50S ribosomal protein L3 2232874:2233584 reverse
c1.0001.2167	K02946	-23,737982	4E-16	-12,601596	2,1E-27	0,0754041	1E-43	c1.0001.2167 30S ribosomal protein S10 2233645:2233953 reverse
c1.0001.2168	K02358	-13,982493	0,0001	-7,3846814	7,9E-12	0,13312364	2E-19	c1.0001.2168 elongation factor Tu 2234410:2235600 reverse
c1.0001.2169	0	-28,581847	1E-79	-15,535409	6,5E-22	0,06289114	1E-28	c1.0001.2169 hypothetical protein, partial 2235638:2236093 reverse
c1.0001.2170	K02355	-8,9846443	4E-30	-4,8215086	8,5E-13	0,19738527	6E-19	c1.0001.2170 elongation factor G, partial 2236036:2237715 reverse
c1.0001.2171	K02992	-10,570284	5E-11	-5,5840566	3,8E-15	0,16981961	5E-23	c1.0001.2171 30S ribosomal protein S7 2237742:2238212 reverse
c1.0001.2172	K02950	-7,9761666	6E-187	-4,367585	1E-06	0,21900195	3E-14	c1.0001.2172 MULTISPECIES: 30S ribosomal protein S12 2238225:2238596 reverse
c1.0001.2174	K03046	-4,2750793	6E-06	-2,2494202	0,00102	0,42307902	2E-08	c1.0001.2174 DNA-directed RNA polymerase subunit beta' 2239756:2243925 reverse
c1.0001.2176	0	-4,2125096	2E-05	-2,1831929	0,00201	0,42317588	2E-04	c1.0001.2176 Hypothetical Protein 2248222:2248461 forward

c1.0001.2177	K02935	-54,254455	5E-06	-28,328498	2,2E-38	0,03364559	1E-53	c1.0001.2177 50S ribosomal protein L7/L12 2248454:2248828 reverse
c1.0001.2178	K02864	-60,66627	1E-22	-32,662573	1,6E-40	0,02994341	3E-58	c1.0001.2178 50S ribosomal protein L10 2248886:2249404 reverse
c1.0001.2179	K02863	-11,923132	1E-56	-6,4235785	2,6E-11	0,14734299	6E-16	c1.0001.2179 50S ribosomal protein L1 2249672:2250373 reverse
c1.0001.2180	K02867	-21,070445	0,0001	-10,985789	1,1E-14	0,08768079	2E-27	c1.0001.2180 50S ribosomal protein L11 2250378:2250812 reverse
c1.0001.2181	K02601	-8,087422	1E-13	-4,2784894	7,9E-10	0,22024839	9E-14	c1.0001.2181 transcription antitermination protein NusG 2250947:2251477 reverse
c1.0001.2182	0	-10,430896	1E-22	-5,6535611	2,9E-08	0,17514015	1E-12	c1.0001.2182 preprotein translocase subunit SecE 2251483:2251680 reverse
c1.0001.2183	K02358	-18,984711	1E-05	-9,8390438	8,6E-20	0,09512056	3E-25	c1.0001.2183 elongation factor Tu 2251979:2253169 reverse
c1.0001.2192	K01997	-9,4015384	0,008	-4,8387436	0,00341	0,20374158	4E-04	c1.0001.2192 ABC transporter permease 2260612:2262453 forward
c1.0001.2195	0	-6,0037686	0,0062	-3,097085	0,00213	0,31162005	9E-07	c1.0001.2195 hypothetical protein 2263929:2264174 reverse
c1.0001.2218	0	-6,7153197	7E-16	-3,7116547	0,01402	0,26609458	0,002	c1.0001.2218 hypothetical protein 2295098:2295358 forward
c1.0001.2226	K02836	-4,0229314	2E-11	-2,1384387	0,00208	0,4415531	4E-05	c1.0001.2226 peptide chain release factor 2 2305574:2306533 reverse
c1.0001.2231	K08300	-4,3549254	5E-05	-2,2967018	0,00012	0,41654565	3E-07	c1.0001.2231 ribonuclease E 2314165:2317338 forward
c1.0001.2233	0	-6,224573	0,0048	-3,1849926	0,00599	0,29704978	7E-04	c1.0001.2233 HAD family hydrolase 2317997:2318608 reverse
c1.0001.2242	0	-8,1539979	1E-07	-4,329195	0,0133	0,22005916	0,002	c1.0001.2242 50S ribosomal protein L11 methyltransferase 2327869:2328528 forward
c1.0001.2244	K03182	-5,0950632	2E-07	-2,778199	0,00195	0,36199584	8E-05	c1.0001.2244 hypothetical protein ABAZ39
c1.0001.2258	K02067	-6,9231426	1E-08	-3,5762636	1,1E-05	0,24996385	1E-06	c1.0001.2258 ATPase AAA 2345091:2345576 forward
c1.0001.2268	0	-4,7544768	4E-05	-2,4546848	0,02413	0,36959491	0,007	c1.0001.2268 hypothetical protein 2355439:2355921 forward
c1.0001.2280	K00762	-8,5884215	2E-06	-4,5143815	0,00029	0,21239541	6E-06	c1.0001.2280 orotate phosphoribosyltransferase 2367348:2368043 forward
c1.0001.2281	K01870	-10,562169	0,001	-5,5421261	1,2E-06	0,17771865	2E-13	c1.0001.2281 isoleucine--tRNA ligase 2368252:2371107 forward
c1.0001.2283	0	-3,7896628	2E-16	-2,0269639	0,01929	0,46102017	0,002	c1.0001.2283 hypothetical protein 2371846:2372451 forward
c1.0001.2287	K01810	-4,1143164	2E-07	-2,1829268	0,00027	0,43743152	5E-07	c1.0001.2287 glucose-6-phosphate isomerase 2375780:2377423 forward
c1.0001.2298	0	-6,2034236	3E-06	-3,3030723	0,03251	0,30137419	0,005	c1.0001.2298 Sodium:solute symporter associated protein 2385350:2385625 forward
c1.0001.2335	K02078	-9,8638863	3E-07	-5,2503496	0,00078	0,19010303	2E-05	c1.0001.2335 acyl carrier protein 2420280:2420540 reverse
c1.0001.2338	0	-7,1272979	2E-06	-3,6249037	0,00194	0,23700592	5E-04	c1.0001.2338 hypothetical protein 2422044:2422445 reverse
c1.0001.2357	0	-6,4722853	3E-06	-3,4674841	0,00508	0,28778635	3E-04	c1.0001.2357 hypothetical protein 2441668:2442843 forward
c1.0001.2358	K08316	-15,654615	0,0085	-7,6266681	0,03282	0,11913685	0,009	c1.0001.2358 DNA methyltransferase 2442834:2443412 reverse
c1.0001.2359	K06178	-4,2157181	2E-09	-2,2443085	0,00984	0,42087777	6E-04	c1.0001.2359 pseudouridine synthase 2443393:2444619 reverse
c1.0001.2366	K01465	-3,8705287	2E-06	-2,0900034	0,04713	0,44254938	0,015	c1.0001.2366 dihydroorotase 2449853:2450911 reverse

c1.0001.2367	K01892	-6,5449669	0,0002	-3,4232938	0,00041	0,28136159	7E-06	c1.0001.2367 histidyl-tRNA synthase 2451067:2452416 forward
c1.0001.2377	K01265	-4,0374439	0,0135	-2,0658841	0,02537	0,45930429	5E-04	c1.0001.2377 methionine aminopeptidase 2462393:2463226 forward
c1.0001.2381	K01756	-9,2482247	5E-14	-5,1153094	3,3E-07	0,20118186	3E-09	c1.0001.2381 adenylosuccinate lyase 2468219:2469514 forward
c1.0001.2384	K01923	-5,6451915	8E-28	-3,0762221	1,8E-05	0,31663761	6E-08	c1.0001.2384 phosphoribosylaminoimidazole-succinocarboxamide synthase 2471029:2471793 forward
c1.0001.2385	K01952	-5,4765644	0,0048	-2,8019839	0,00219	0,33601363	3E-05	c1.0001.2385 phosphoribosylformylglycinamidine synthase 2471856:2472098 forward
c1.0001.2386	K01247	-7,9457922	0,0002	-4,1230441	0,01895	0,23232864	0,004	c1.0001.2386 DNA glycosylase 2472245:2472898 forward
c1.0001.2390	K01952	-4,4940496	2E-13	-2,4264958	0,00509	0,40099631	1E-04	c1.0001.2390 phosphoribosylformylglycinamidine synthase 2475877:2476581 forward
c1.0001.2391	K01952	-7,4461484	6E-12	-3,9694123	4,6E-10	0,24100548	2E-15	c1.0001.2391 phosphoribosylformylglycinamidine synthase 2476578:2478794 forward
c1.0001.2392	0	-4,6255065	0,0007	-2,4226157	5,1E-05	0,39600752	4E-07	c1.0001.2392 ATP-binding protein 2479093:2479332 forward
c1.0001.2393	K07390	-7,3580522	3E-64	-3,9557842	1,2E-10	0,24011039	6E-16	c1.0001.2393 glutaredoxin 2479394:2479735 forward
c1.0001.2402	K02533	-14,801509	0,0001	-7,6645723	8,5E-05	0,12533797	2E-06	c1.0001.2402 RNA methyltransferase 2486274:2487044 reverse
c1.0001.2403	K01649	-12,196668	4E-07	-6,4433018	3,6E-13	0,14954309	2E-18	c1.0001.2403 2-isopropylmalate synthase 2487061:2488665 reverse
c1.0001.2406	0	-7,3190277	1E-14	-3,9636471	9E-05	0,25038023	3E-07	c1.0001.2406 preQ0 transporter 2491405:2491869 reverse
c1.0001.2408	K01883	-4,0200239	1E-06	-2,2025495	0,02952	0,45310382	0,004	c1.0001.2408 cysteinyl-tRNA synthetase 2493386:2494759 reverse
c1.0001.2422	0	-15,017158	4E-08	-7,98631	0,00355	0,12291415	5E-04	c1.0001.2422 hypothetical protein 2511069:2511707 forward
c1.0001.2423	K01999	-17,526923	0,0007	-9,1137947	6,4E-08	0,10792123	2E-11	c1.0001.2423 ABC transporter substrate-binding protein 2511759:2512925 forward
c1.0001.2432	K02888	-10,33873	5E-13	-5,5768329	1,3E-13	0,17516603	9E-17	c1.0001.2432 50S ribosomal protein L21 2521821:2522138 forward
c1.0001.2434	K03979	-7,3574209	8E-10	-3,8454873	1,8E-06	0,23791875	6E-09	c1.0001.2434 GTPase CgtA 2522597:2523820 forward
c1.0001.2435	K00931	-5,1134956	5E-05	-2,8259132	0,03535	0,34290153	0,012	c1.0001.2435 glutamate 5-kinase 2523765:2524961 forward
c1.0001.2441	K00147	-5,0656547	9E-05	-2,6464317	0,00116	0,35780127	2E-05	c1.0001.2441 gamma-glutamyl phosphate reductase 2529471:2530751 forward
c1.0001.2451	K02114	-11,373595	9E-05	-5,9748918	2,9E-11	0,16322894	3E-15	c1.0001.2451 ATP synthase subunit epsilon 2539776:2540189 reverse
c1.0001.2452	K02112	-8,5742375	0,001	-4,4204929	5,7E-07	0,21468659	3E-16	c1.0001.2452 F0F1 ATP synthase subunit beta 2540280:2541719 reverse
c1.0001.2453	K02115	-6,5089498	0,0004	-3,3780953	6E-08	0,28122318	2E-12	c1.0001.2453 ATP synthase F0F1 subunit gamma 2541789:2542679 reverse
c1.0001.2454	K02111	-3,9430231	0,0001	-2,0626883	0,00096	0,46125618	1E-06	c1.0001.2454 ATP F0F1 synthase subunit alpha 2542752:2544281 reverse

c1.0001.2459	K00616	-4,0427959	0,003	-2,0993241	0,00153	0,45481213	1E-05	c1.0001.2459 transaldolase 2548691:2549347 forward
c1.0001.2462	K00382	-6,1308291	8E-11	-3,2590528	2E-06	0,29211188	2E-09	c1.0001.2462 dihydrolipoamide dehydrogenase 2551124:2552524 reverse
c1.0001.2469	0	-6,6557481	7E-22	-3,645084	0,00016	0,27464701	1E-06	c1.0001.2469 hypothetical protein ABAZ39
c1.0001.2475	K00927	-10,678241	1E-14	-5,788838	8,3E-11	0,17134543	9E-15	c1.0001.2475 phosphoglycerate kinase 2563508:2564704 forward
c1.0001.2477	K09747	-3,6561195	6E-30	-2,0073884	0,01408	0,47884251	0,001	c1.0001.2477 hypothetical protein 2565521:2565844 reverse
c1.0001.2485	0	-9,7057624	8E-11	-5,0034352	1,8E-06	0,17620749	3E-08	c1.0001.2485 hypothetical protein 2573077:2573190 reverse
c1.0001.2492	K13896	-4,1949674	1E-08	-2,2560317	0,00985	0,42740706	5E-04	c1.0001.2492 microcin ABC transporter ATP-binding protein 2581515:2583161 forward
c1.0001.2540	K03704	-3,6847678	5E-40	-2,0096816	0,0049	0,4787427	3E-04	c1.0001.2540 cold-shock protein 2629658:2629867 forward
c1.0001.2541	0	-5,2513993	0,0004	-2,7664038	0,03754	0,35827376	0,006	c1.0001.2541 hypothetical protein 2629976:2630206 forward
c1.0001.2549	K01703	-21,590254	2E-07	-11,307703	5E-18	0,08365766	3E-24	c1.0001.2549 isopropylmalate isomerase 2635695:2637098 reverse
c1.0001.2553	K02884	-17,284143	1E-24	-9,4436078	2,2E-16	0,1065029	7E-21	c1.0001.2553 50S ribosomal protein L19 2638780:2639175 reverse
c1.0001.2555	K02860	-23,714089	1E-119	-12,955276	3E-14	0,07464606	5E-26	c1.0001.2555 16S rRNA-processing protein RimM 2640013:2640618 reverse
c1.0001.2556	K02959	-15,191598	2E-07	-8,1856303	1,2E-14	0,12326961	7E-18	c1.0001.2556 30S ribosomal protein S16 2640715:2641068 reverse
c1.0001.2557	K03106	-9,2189035	5E-22	-4,9656946	1,1E-08	0,19377953	5E-13	c1.0001.2557 signal recognition particle 2641150:2642517 reverse
c1.0001.2561	0	-10,157475	3E-05	-5,284594	9,1E-12	0,17868347	9E-16	c1.0001.2561 hypothetical protein 2645786:2647120 reverse
c1.0001.2564	K00674	-8,8191559	2E-06	-4,5803042	1E-06	0,20205021	1E-09	c1.0001.2564 2,3,4,5-tetrahydropyridine-2,6-carboxylate N-succinyltransferase 2649815:2650657 forward
c1.0001.2565	K01439	-5,0145771	1E-07	-2,73459	0,01166	0,36666607	9E-04	c1.0001.2565 succinyl-diaminopimelate desuccinylase 2650677:2651858 forward
c1.0001.2566	0	-9,0893667	5E-13	-4,9583823	0,00129	0,20015375	4E-05	c1.0001.2566 membrane protein 2651906:2652475 forward
c1.0001.2567	K00768	-8,6589741	0,0227	-4,1926121	0,04346	0,20248846	0,004	c1.0001.2567 nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase 2652587:2653615 reverse
c1.0001.2572	K02536	-5,4619781	6E-07	-2,9666834	0,0096	0,3152221	0,002	c1.0001.2572 UDP-3-O-(3-hydroxymyristoyl) glucosamine N-acyltransferase 2655873:2657000 reverse
c1.0001.2573	K01992	-11,2945	6E-07	-6,0960424	0,0058	0,16633594	5E-04	c1.0001.2573 ABC transporter 2657081:2657893 reverse
c1.0001.2597	K03217	-5,6578335	6E-28	-3,0992248	1,2E-05	0,31100834	2E-08	c1.0001.2597 insertase 2683244:2684983 forward
c1.0001.2598	K03978	-7,9818152	0,0011	-4,0584269	0,03615	0,22709136	0,012	c1.0001.2598 GTP-binding protein YsxC 2685128:2685790 forward
c1.0001.2601	K00930	-5,0424419	0,0041	-2,6451553	0,00797	0,37591778	5E-04	c1.0001.2601 acetylglutamate kinase 2687499:2688398 forward
c1.0001.2606	K01962	-11,23332	5E-06	-5,9622922	8,1E-08	0,16627593	8E-12	c1.0001.2606 acetyl-CoA carboxylase subunit alpha 2690653:2691609 reverse
c1.0001.2609	0	-5,9558937	1E-285	-3,208865	3,3E-07	0,29437074	7E-11	c1.0001.2609 Hypothetical Protein 2694607:2694966 reverse
c1.0001.2610	0	-6,8372234	0,0029	-3,4804654	0,0002	0,27095679	2E-09	c1.0001.2610 Hypothetical Protein 2694971:2695420 reverse
c1.0001.2616	K00891	-10,960615	0,0001	-5,6475555	1,6E-06	0,16498123	6E-09	c1.0001.2616 shikimate kinase 2699065:2699700 forward

c1.0001.2627	K02012	-5,9358164	2E-05	-3,0944959	9,2E-05	0,30322051	4E-07	c1.0001.2627 iron deficiency-induced protein A 2708441:2709469 forward
c1.0001.2635	0	#DIV/0!	3E-16	#DIV/0!	0,00119	0,01012422	2E-05	c1.0001.2635 hypothetical protein 2716342:2716485 reverse
c1.0001.2641	K02945	-5,1225378	3E-10	-2,724534	4,6E-06	0,35004969	3E-11	c1.0001.2641 30S ribosomal protein S1 2720736:2722433 reverse
c1.0001.2644	K00800	-8,1057965	2E-05	-4,1522019	0,00368	0,21058968	8E-04	c1.0001.2644 3-phosphoshikimate 1-carboxyvinyltransferase 2723269:2724606 reverse
c1.0001.2658	0	-3,8953441	6E-10	-2,0619928	0,00234	0,4579291	9E-05	c1.0001.2658 heat shock protein Hs
c1.0001.2667	0	-4,1414721	3E-09	-2,2375484	0,01942	0,42270019	0,003	c1.0001.2667 hypothetical protein 2752451:2753470 reverse
c1.0001.2684	K09792	-5,0266738	0,0027	-2,6025229	0,00467	0,36829034	3E-04	c1.0001.2684 membrane protein 2768169:2768936 reverse
c1.0001.2687	0	-6,8739246	4E-06	-3,6579617	0,01671	0,27235421	0,003	c1.0001.2687 ankyrin 2769969:2770367 forward
c1.0001.2690	0	-12,626975	7E-05	-6,6289967	1,2E-14	0,14527052	3E-23	c1.0001.2690 RNA helicase 2771843:2773372 reverse
c1.0001.2692	K03070	-6,3423482	4E-11	-3,3761491	8,4E-08	0,28218177	2E-11	c1.0001.2692 preprotein translocase subunit SecA 2774238:2776988 reverse
c1.0001.2693	K03769	-7,6751103	7E-11	-4,0913866	0,00054	0,22595047	3E-05	c1.0001.2693 peptidylprolyl isomerase 2777375:2778322 forward
c1.0001.2701	0	-63,232777	0,0001	-34,501854	2,3E-06	0,03618676	1E-08	c1.0001.2701 DMT transporter permease 2785313:2786227 forward
c1.0001.2703	0	-34,701974	0,0003	-18,881048	0,00109	0,06590817	3E-05	c1.0001.2703 C4-dicarboxylate ABC transporter permease 2787514:2787996 reverse
c1.0001.2704	0	-4,4125122	0,0002	-2,3175754	0,00063	0,41429051	1E-05	c1.0001.2704 C4-dicarboxylate ABC transporter 2788113:2789141 reverse
c1.0001.2733	0	-6,6327014	3E-10	-3,6166841	0,00012	0,27412979	6E-07	c1.0001.2733 signal peptide protein 2815243:2817138 forward
c1.0001.2754	K04077	-25,348454	9E-15	-13,441615	2E-27	0,07066038	4E-49	c1.0001.2754 molecular chaperone GroEL 2839598:2841241 reverse
c1.0001.2755	K04078	-31,596502	0,0001	-16,68263	4,2E-15	0,06004946	1E-28	c1.0001.2755 molecular chaperone GroES 2841294:2841584 reverse
c1.0001.2786	K00769	-7,6803527	2E-26	-4,1168854	4E-05	0,22956667	2E-07	c1.0001.2786 xanthine phosphoribosyltransferase 2877877:2878377 forward
c1.0001.2812	0	-3,8616479	3E-08	-2,0305267	0,00094	0,46682347	6E-06	c1.0001.2812 Hypothetical Protein 2902170:2903438 forward
c1.0001.2836	0	-6,6755895	0,0122	-3,478663	0,01628	0,29103246	4E-04	c1.0001.2836 Hypothetical Protein 2923080:2924099 reverse
c1.0001.2855	0	-8,5005318	0,0001	-4,4434158	7,6E-06	0,21683215	1E-08	c1.0001.2855 Hypothetical Protein 2942927:2944300 reverse
c1.0001.2856	0	-6,2219324	0,0053	-3,2086954	0,00101	0,29935523	9E-08	c1.0001.2856 Hypothetical Protein 2944297:2944950 reverse
c1.0001.2857	K07337	-6,5708322	3E-09	-3,4764959	0,02103	0,26436558	0,006	c1.0001.2857 hypothetical protein 2945267:2945743 forward
c1.0001.2879	K01975	-4,4830918	0,0003	-2,3301497	0,02929	0,40358645	0,006	c1.0001.2879 2'-5' RNA ligase 2967510:2968046 reverse
c1.0001.2886	K03686	-4,2339871	2E-17	-2,2705734	0,00012	0,42049687	3E-07	c1.0001.2886 molecular chaperone DnaJ 2975922:2977064 reverse
c1.0001.2887	K04043	-4,8799448	1E-05	-2,5598504	3,5E-05	0,37159335	3E-10	c1.0001.2887 molecular chaperone DnaK 2977226:2979145 reverse
c1.0001.2889	0	-6,6394265	3E-15	-3,6749735	0,00122	0,2747322	6E-05	c1.0001.2889 hypothetical protein ABAZ39

c1.0001.2893	K03687	-11,75586	7E-14	-6,2188822	1,3E-11	0,15147708	2E-16	c1.0001.2893 heat shock protein GrpE 2983987:2984595 reverse
c1.0001.2894	K03705	-12,38114	0,0092	-6,3192196	0,00093	0,15371598	1E-06	c1.0001.2894 heat-inducible transcription repressor 2984675:2985715 reverse
c1.0001.2898	K03406	-6,6325134	0,0003	-3,4680221	9,3E-05	0,27749818	4E-07	c1.0001.2898 chemotaxis protein 2988693:2990384 forward
c1.0001.2899	K00287	-18,902163	5E-05	-10,39308	0,03559	0,1187448	0,01	c1.0001.2899 dihydrofolate reductase 2990418:2990924 forward
P1.0002.0012	K01923	-6,2912205	3E-07	-3,3207247	5,8E-06	0,28746493	3E-09	P1.0002.0012 phosphoribosylaminoimidazole-succinocarboxamide synthase 3015215:3016198 forward
P1.0002.0014	0	-5,1258807	0,0129	-2,6626877	0,02602	0,37020199	8E-05	P1.0002.0014 nitrogen regulatory protein 3017619:3018107 forward
P1.0002.0097	K04758	-6,6776922	5E-24	-3,7181064	0,00067	0,27063587	1E-06	P1.0002.0097 iron transporter FeoA 3119439:3119738 forward
P1.0002.0098	K04759	-13,574637	5E-05	-7,0440748	4E-14	0,13387959	1E-18	P1.0002.0098 iron transporter FeoB 3119806:3121695 forward
P1.0002.0114	0	-5,3543756	0,0013	-2,7836803	0,0273	0,34481156	0,006	P1.0002.0114 metallo-beta-lactamase 3136085:3137083 reverse
P1.0002.0122	0	-8,0701591	1E-05	-4,2272265	3E-11	0,22537657	2E-18	P1.0002.0122 tyrosyl-tRNA synthetase, partial 3144912:3145997 reverse
P1.0002.0167	K05837	-4,542248	0,0107	-2,2925129	0,0481	0,39768473	0,01	P1.0002.0167 cell wall shape-determining protein 3210234:3211406 reverse
P1.0002.0171	K01586	-3,9725356	3E-11	-2,1317194	0,00772	0,4494785	2E-04	P1.0002.0171 diaminopimelate decarboxylase 3215790:3217055 reverse
P1.0002.0176	K00074	-11,122319	2E-20	-5,9707574	1E-14	0,16125886	4E-21	P1.0002.0176 3-hydroxybutyryl-CoA dehydrogenase 3220252:3221130 reverse
P1.0002.0177	K03522	-10,426366	4E-06	-5,5450023	3,4E-11	0,17743006	5E-15	P1.0002.0177 electron transfer flavoprotein subunit beta 3221185:3222123 reverse
P1.0002.0178	K03521	-12,177253	1E-90	-6,6106957	3,9E-11	0,14373347	1E-18	P1.0002.0178 electron transfer flavoprotein subunit beta 3222159:3222923 reverse
P1.0002.0193	0	-4,440348	3E-08	-2,4092202	0,02985	0,40648014	0,004	P1.0002.0193 hypothetical protein ABAZ39
P1.0002.0198	K07091	-4,3400305	2E-12	-2,3600558	0,00661	0,41291739	3E-04	P1.0002.0198 permease 3240046:3241167 forward
P1.0002.0199	0	-12,674625	0,0052	-6,5709097	0,00362	0,15568901	3E-04	P1.0002.0199 glycosyl transferase 3241185:3242198 forward
P1.0002.0207	0	-7,3945152	8E-05	-3,8581294	6,9E-07	0,2460197	2E-10	P1.0002.0207 histidine kinase 3252279:3254750 reverse
P1.0002.0209	K00375	-6,0188883	0,0002	-3,3248041	0,03233	0,31543868	0,005	P1.0002.0209 GntR family transcriptional regulator 3256719:3258161 forward
P1.0002.0226	K02234	-4,5559128	3E-09	-2,5078984	0,0132	0,37726008	0,002	P1.0002.0226 cobalamin biosynthesis protein CobW 3274594:3275667 reverse
P1.0002.0230	0	-21,869379	0,0474	-10,393652	0,03967	0,08811725	0,003	P1.0002.0230 histidinol dehydrogenase 3277698:3277835 forward
P1.0002.0237	K11209	-4,5195246	3E-42	-2,4672433	0,00166	0,38333543	8E-06	P1.0002.0237 glutathione S-transferase 3284063:3284764 reverse
P1.0002.0262	K07112	-22,375545	3E-07	-12,068835	9,1E-05	0,08132647	6E-06	P1.0002.0262 membrane protein 3312329:3313564 forward
P1.0002.0320	0	-4,9552091	6E-06	-2,5832088	6E-05	0,36222401	2E-06	P1.0002.0320 Hypothetical Protein 3364761:3365579 forward

P1.0002.0324	K03427	-7,7064062	0,0008	-3,963054	5,6E-05	0,23557495	7E-07	P1.0002.0324 N-6 DNA methylase:Type I restriction-modification system, M subunit 3368728:3370470 forward
P1.0002.0325	0	-9,0788389	0,0004	-4,7034508	0,0315	0,18038677	0,014	P1.0002.0325 Hypothetical Protein 3370467:3371402 forward
P1.0002.0342	0	-5,7228601	3E-09	-2,9998718	8,4E-08	0,31430809	8E-10	P1.0002.0342 Hypothetical Protein 3385087:3385887 reverse
P1.0002.0343	0	-5,3088122	0,0005	-2,731296	4,1E-05	0,34529575	2E-07	P1.0002.0343 Hypothetical Protein 3386242:3387060 reverse
P1.0002.0345	0	-18,569479	1E-05	-9,275341	0,01224	0,09743855	0,002	P1.0002.0345 Hypothetical Protein 3387545:3387739 forward
P1.0002.0357	K00411	-8,2853991	9E-13	-4,4596592	1,8E-06	0,22076317	5E-10	P1.0002.0357 ubiquinol-cytochrome C reductase 3401373:3401942 forward
P1.0002.0358	K00412	-6,0057996	7E-08	-3,1995366	1,7E-07	0,30121959	7E-11	P1.0002.0358 cytochrome B 3401972:3403240 forward
P1.0002.0359	K00413	-10,041596	3E-66	-5,3914228	1,7E-11	0,17337565	1E-14	P1.0002.0359 cytochrome C 3403240:3404004 forward
P1.0002.0366	0	-8,2453066	5E-15	-4,4286661	3,8E-11	0,21803613	5E-16	P1.0002.0366 acetoin utilization protein 3406837:3407766 forward
P1.0002.0369	K01662	-4,0338257	1E-06	-2,1538114	0,00383	0,45018438	8E-05	P1.0002.0369 1-deoxy-D-xylulose-5-phosphate synthase 3409088:3411004 forward
P1.0002.0380	0	-6,3828166	1E-09	-3,4628784	0,04737	0,27338536	0,02	P1.0002.0380 Hypothetical Protein 3420512:3420826 forward
P1.0002.0429	0	-3,9288586	3E-15	-2,1210694	0,0076	0,45701375	2E-04	P1.0002.0429 hypothetical protein 3465006:3465260 forward
P1.0002.0441	K01255	-4,9574927	2E-07	-2,6034318	0,00014	0,35952357	1E-06	P1.0002.0441 aminopeptidase A 3475170:3476672 reverse
P1.0002.0442	K11720	-4,1625218	1E-06	-2,2330574	0,01498	0,44006052	8E-04	P1.0002.0442 permease 3476899:3477993 forward
P1.0002.0443	K03771	-5,9868277	1E-06	-3,1798538	0,00088	0,30405828	1E-05	P1.0002.0443 peptidylprolyl isomerase 3478133:3479572 forward
P1.0002.0455	0	-5,3930435	0,0009	-2,7995037	0,01488	0,34123071	0,002	P1.0002.0455 phospholipid N-methyltransferase 3489830:3490444 reverse
P1.0002.0456	0	-12,033955	4E-06	-6,4610217	0,01367	0,15131559	0,002	P1.0002.0456 hypothetical protein 3490588:3491289 reverse
P1.0002.0457	K06158	-6,174611	3E-09	-3,2567275	4,7E-06	0,28784068	6E-09	P1.0002.0457 glycosyl transferase family 1 3491286:3493181 reverse
P1.0002.0458	K00940	-17,960388	3E-51	-9,8798483	2,9E-09	0,10001102	1E-13	P1.0002.0458 MULTISPECIES: nucleoside diphosphate kinase 3493384:3493806 forward
P1.0002.0462	K01933	-6,7040093	6E-09	-3,5712249	0,00114	0,26917538	5E-05	P1.0002.0462 phosphoribosylformylglycinamidine cyclo-ligase 3495405:3496439 reverse
P1.0002.0473	K03684	-3,842616	1E-05	-2,0331282	0,01379	0,47224688	7E-04	P1.0002.0473 ribonuclease D 3509147:3510313 reverse
P1.0002.0474	K01876	-4,7303005	9E-06	-2,4796997	4,4E-05	0,38165243	4E-07	P1.0002.0474 aspartyl-tRNA synthetase 3510480:3512309 forward
P1.0002.0483	K02913	-4,780132	0,0009	-2,4873684	2,2E-05	0,38335528	3E-08	P1.0002.0483 MULTISPECIES: 50S ribosomal protein L33 3519758:3519925 reverse
P1.0002.0498	0	-10,106074	8E-09	-5,8556715	0,01576	0,18973698	8E-04	P1.0002.0498 polyketide cyclase 3532995:3533468 reverse
P1.0002.0500	0	-5,412227	0,003	-2,7738512	0,00757	0,33654554	0,001	P1.0002.0500 hypothetical protein 3534330:3535565 reverse
P1.0002.0559	K21395	-14,658498	0,0003	-8,0154674	0,00196	0,14251036	1E-04	P1.0002.0559 ABC transporter substrate-binding protein 3592229:3593263 forward
P1.0002.0564	K03168	-5,6051492	8E-10	-2,9619883	6,7E-06	0,31627396	4E-08	P1.0002.0564 DNA topoisomerase I 3596627:3599491 reverse

P1.0002.0577	K07772	-5,6175912	2E-06	-2,997498	0,00206	0,32912229	5E-05	P1.0002.0577 response regulator receiver 3614390:3615139 forward
P1.0002.0589	K01772	-13,396167	2E-07	-7,0306232	1,1E-05	0,1349876	4E-08	P1.0002.0589 ferrochelatase 3626373:3627512 forward
P1.0002.0599	K00033	-4,0973898	0,0003	-2,156967	0,00136	0,44699269	4E-05	P1.0002.0599 6-phosphogluconate dehydrogenase 3635558:3636997 reverse
P1.0002.0600	K06145	-8,1288762	2E-12	-4,5435501	0,00115	0,2196626	8E-06	P1.0002.0600 LacI family transcription regulator 3637098:3638105 reverse
P1.0002.0603	0	-4,6383988	0,0091	-2,398574	0,01142	0,40494571	0,001	P1.0002.0603 2-keto-4-pentenoate hydratase 3640850:3641614 reverse
P1.0002.0611	K09948	-3,8774084	7E-15	-2,0826495	0,00764	0,46173933	2E-04	P1.0002.0611 alkaline phosphatase 3648401:3648727 reverse
P1.0002.0612	K09138	-5,718166	5E-05	-3,0303531	0,02965	0,31791579	0,006	P1.0002.0612 hypothetical protein 3648739:3649977 reverse
P1.0002.0618	K01494	-8,3019558	1E-05	-4,4058913	7,6E-07	0,22403486	2E-10	P1.0002.0618 deoxycytidine triphosphate deaminase 3653810:3654394 forward
P1.0002.0636	K01129	-6,0361868	4E-17	-3,199351	3,3E-06	0,29076572	1E-08	P1.0002.0636 deoxyguanosinetriphosphate triphosphohydrolase 3672518:3673717 forward
P1.0002.0637	K01887	-7,2330598	1E-13	-3,8788444	5,9E-06	0,24250978	7E-09	P1.0002.0637 arginyl-tRNA synthetase 3673779:3675539 forward
P1.0002.0640	0	-8,6849131	7E-08	-4,7033659	0,00722	0,2018713	0,001	P1.0002.0640 peptidase M48 3677604:3678290 forward
P1.0002.0644	K03116	-4,8173597	7E-45	-2,6043759	0,00079	0,3601572	2E-05	P1.0002.0644 preprotein translocase subunit TatA 3681301:3681594 forward
P1.0002.0645	K03117	-10,39621	2E-32	-5,5096662	2,6E-08	0,16614489	1E-11	P1.0002.0645 preprotein translocase subunit TatB 3681643:3682164 forward
P1.0002.0648	K01875	-5,9668102	0,0088	-3,0470245	0,00361	0,31085566	2E-05	P1.0002.0648 serine--tRNA ligase 3684353:3685618 forward
P1.0002.0649	K03787	-8,5335388	6E-06	-4,4793363	0,00032	0,21411247	6E-06	P1.0002.0649 stationary phase survival protein SurE 3685618:3686400 forward
P1.0002.0652	K06923	-5,7345185	1E-06	-3,0466204	0,00027	0,31934266	1E-06	P1.0002.0652 ATPase AAA 3688343:3689200 reverse
P1.0002.0653	0	-9,5385378	8E-05	-5,2455567	0,0389	0,185045	0,014	P1.0002.0653 hypothetical protein 3689339:3690079 forward
P1.0002.0654	K03210	-9,7226756	0,0488	-4,9668127	0,04374	0,20251138	2E-06	P1.0002.0654 preprotein translocase subunit YajC 3690191:3690562 forward
P1.0002.0655	K03072	-9,881195	0,0013	-5,109353	1,7E-06	0,1875855	5E-11	P1.0002.0655 preprotein translocase subunit SecD 3690599:3692197 forward
P1.0002.0656	K03074	-9,3884645	2E-16	-5,0491838	2,9E-08	0,19250535	9E-13	P1.0002.0656 preprotein translocase subunit SecF 3692210:3693163 forward
P1.0002.0657	0	-10,460239	4E-11	-5,7134658	0,00888	0,1768704	0,001	P1.0002.0657 valyl-tRNA synthetase 3693206:3693583 forward
P1.0002.0658	K04564	-5,3911633	5E-20	-2,8943994	7,6E-07	0,33049202	1E-10	P1.0002.0658 superoxide dismutase 3693669:3694274 reverse
P1.0002.0685	K01687	-4,9730641	0,0003	-2,5620424	0,00013	0,36395532	4E-06	P1.0002.0685 dihydroxy-acid dehydratase 3721925:3723790 reverse

P1.0002.0703	0	-4,1133148	4E-06	-2,1653093	0,00029	0,43937554	2E-06	P1.0002.0703 hypothetical protein 3745295:3745837 reverse
P1.0002.0711	K00950	-5,565726	0,0083	-2,8202324	0,00983	0,32884941	0,001	P1.0002.0711 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase 3751382:3751888 reverse
P1.0002.0742	K00133	-6,6720999	2E-19	-3,541413	5,3E-09	0,26745431	4E-15	P1.0002.0742 aspartate-semialdehyde dehydrogenase 3785482:3786495 forward
P1.0002.0743	0	-9,4162718	6E-05	-4,9200568	0,00048	0,19679431	2E-05	P1.0002.0743 Hypothetical Protein 3786702:3787187 forward
P1.0002.0744	K02837	-5,3608075	1E-11	-2,8530697	3,5E-06	0,33354113	5E-09	P1.0002.0744 peptide chain release factor 3 3787345:3788952 forward
P1.0002.0768	0	-9,0557662	2E-19	-4,8233438	6,2E-10	0,19714751	1E-24	P1.0002.0768 ABC transporter substrate-binding protein 3818093:3819199 forward
P1.0002.0785	K12140	-4,311995	1E-09	-2,2684457	0,01677	0,403765	0,004	P1.0002.0785 hydrogenase-4 component E 3836137:3836799 forward
P1.0002.0786	K12141	-5,2338445	0,0003	-2,7588365	0,00109	0,35397771	1E-05	P1.0002.0786 oxidoreductase 3836796:3838253 forward
P1.0002.0789	K04765	-7,2663389	9E-07	-3,8853256	0,01074	0,25623113	0,001	P1.0002.0789 nucleoside triphosphate hydrolase 3840346:3841155 reverse
P1.0002.0796	K03666	-9,2639228	6E-29	-4,8924174	8,4E-12	0,18917568	7E-14	P1.0002.0796 RNA-binding protein Hfq 3845433:3845681 reverse
P1.0002.0800	0	-4,2327383	2E-17	-2,2946254	0,03461	0,3986062	0,005	P1.0002.0800 hypothetical protein 3848978:3849286 reverse
P1.0002.0810	K03644	-4,8906969	7E-08	-2,5977634	0,00017	0,36960268	4E-07	P1.0002.0810 radical SAM protein 3860502:3861458 reverse
P1.0002.0811	K00382	-5,3266928	2E-09	-2,8255627	2,9E-06	0,33716971	3E-12	P1.0002.0811 dihydrolipoamide dehydrogenase 3861721:3863118 reverse
P1.0002.0812	K00627	-12,304934	0,0017	-6,3769162	5,7E-07	0,15044754	1E-21	P1.0002.0812 branched-chain alpha-keto acid dehydrogenase subunit E2 3863131:3863814 reverse
P1.0002.0813	K00162	-14,037657	5E-10	-7,4283921	1,3E-16	0,12827618	4E-35	P1.0002.0813 pyruvate dehydrogenase E1 component, beta subunit 3864485:3865882 reverse
P1.0002.0814	K00161	-5,3363327	1E-06	-2,8170854	2,8E-06	0,33860537	1E-11	P1.0002.0814 pyruvate dehydrogenase E1 subunit alpha 3865961:3866977 reverse
P1.0002.0821	0	-6,1045838	1E-26	-3,3198117	1,5E-06	0,29041684	3E-09	P1.0002.0821 DNA helicase 3872515:3874077 forward
P1.0002.0822	0	-17,966354	8E-21	-10,410083	0,00093	0,10862487	6E-06	P1.0002.0822 AsnC family transcriptional regulator 3874083:3874325 forward
P1.0002.0829	K01802	-10,750652	1E-40	-5,8441933	6,7E-08	0,1659241	9E-12	P1.0002.0829 peptidylprolyl isomerase 3882023:3882574 forward
P1.0002.0844	K01358	-3,6650725	1E-75	-2,0060198	0,00634	0,4743964	8E-05	P1.0002.0844 ATP-dependent Clp protease ClpP 3904907:3905533 forward
P1.0002.0853	K03100	-13,52143	4E-13	-7,6800182	0,00167	0,1349913	2E-06	P1.0002.0853 signal peptidase 3912553:3913323 reverse
P1.0002.0857	K00951	-4,2358913	0,0002	-2,2030828	0,00028	0,42845996	1E-05	P1.0002.0857 GTP pyrophosphokinase 3915073:3917244 reverse
P1.0002.0863	0	-4,7756388	0,0099	-2,4740484	0,01589	0,39484267	3E-04	P1.0002.0863 transcriptional regulator 3921156:3921689 reverse

P1.0002.0868	0	-19,612402	0,0066	-10,517823	0,03487	0,11792957	0,009	P1.0002.0868 conserved exported protein of unknown function 3926022:3926273 reverse
P1.0002.0872	0	-4,2265524	1E-04	-2,2010679	0,00545	0,42309194	4E-04	P1.0002.0872 protein phosphatase 3930152:3932125 reverse
P1.0002.0878	0	-5,1921346	1E-09	-2,9160509	0,04844	0,341243	0,014	P1.0002.0878 hypothetical protein 3936449:3936838 forward
P1.0002.0885	0	-5,2663864	7E-24	-2,825331	1,1E-06	0,33765915	2E-10	P1.0002.0885 lytic transglycosylase 3942366:3943031 forward
P1.0002.0886	K07115	-6,8183	0,0151	-3,5105422	0,01074	0,27987619	8E-06	P1.0002.0886 lactate dehydrogenase 3943089:3943928 forward
P1.0002.0905	K09767	-6,9328016	0,0002	-3,6830227	0,00048	0,27372205	6E-06	P1.0002.0905 nucleotide-binding protein 3964845:3965333 reverse
P1.0002.0909	K07216	-4,1535514	3E-63	-2,264936	0,0006	0,42129535	3E-05	P1.0002.0909 hemerythrin 3967728:3968138 reverse
P1.0002.0911	K12372	-5,1303936	3E-05	-2,7003911	0,04357	0,3416895	0,011	P1.0002.0911 peptide ABC transporter ATP-binding protein 3968621:3969625 reverse
P1.0002.0914	K12369	-3,832117	1E-05	-2,0219122	0,03901	0,46887103	0,006	P1.0002.0914 peptide ABC transporter permease 3971558:3972568 reverse
P1.0002.0933	K01433	-5,5430167	6E-05	-2,923032	0,00761	0,33239543	4E-04	P1.0002.0933 formyltetrahydrofolate deformylase 3989743:3990609 reverse
P1.0002.0975	K00759	-12,844923	6E-07	-7,5609164	0,03866	0,15129404	0,003	P1.0002.0975 adenine phosphoribosyltransferase 4036254:4036781 forward
P1.0002.1002	K01951	-4,0054899	0,0029	-2,0922562	0,00614	0,46251985	3E-04	P1.0002.1002 GMP synthase 4066633:4068183 reverse
P1.0002.1020	0	-11,048839	3E-24	-6,1006353	2,4E-05	0,15970591	8E-09	P1.0002.1020 GntR family transcriptional regulator 4086135:4086836 reverse
P1.0002.1021	0	-5,3264226	2E-08	-2,7877341	0,00843	0,32464265	0,001	P1.0002.1021 phytanoyl-CoA dioxygenase 4086901:4087809 reverse
P1.0002.1022	0	-10,950161	4E-05	-5,7753348	0,00368	0,15847299	7E-04	P1.0002.1022 C4-dicarboxylate ABC transporter permease 4087826:4089145 reverse
P1.0002.1088	K11103	-5,1208626	6E-09	-2,7239579	4,6E-06	0,35075782	9E-11	P1.0002.1088 C4-dicarboxylate ABC transporter 4160240:4161541 reverse
P1.0002.1089	K10125	-4,2782293	9E-05	-2,3132337	0,0436	0,42068629	0,007	P1.0002.1089 histidine kinase 4161878:4163599 forward
P1.0002.1090	K10126	-4,5597313	7E-07	-2,4348878	0,00714	0,39899437	3E-04	P1.0002.1090 Fis family transcriptional regulator 4163596:4164975 forward
P1.0002.1096	K01965	-5,6316044	9E-16	-3,0246304	2E-06	0,31693416	2E-09	P1.0002.1096 acetyl-CoA carboxylase 4171006:4173015 reverse
P1.0002.1097	K01966	-4,1692762	1E-13	-2,2322403	0,00119	0,4257315	2E-05	P1.0002.1097 methylmalonyl-CoA carboxyltransferase 4173030:4174562 reverse
P1.0002.1105	K02434	-7,6926892	2E-15	-4,1206596	5,1E-08	0,23278275	9E-12	P1.0002.1105 glutamyl-tRNA amidotransferase 4180790:4182241 reverse
P1.0002.1106	K02433	-5,3106717	1E-13	-2,891583	0,00036	0,33517137	3E-06	P1.0002.1106 glutamyl-tRNA amidotransferase 4182263:4183744 reverse

P1.0002.1107	K02435	-7,9995377	4E-30	-4,3670759	0,00027	0,22085178	8E-06	P1.0002.1107 glutamyl-tRNA amidotransferase subunit C 4183744:4184031 reverse
P1.0002.1110	K07088	-24,383177	1E-09	-12,974477	4,3E-05	0,07587048	7E-07	P1.0002.1110 transporter 4185557:4186486 forward
P1.0002.1111	0	-93,007407	0,0023	-47,765657	1,7E-12	0,0200837	6E-56	P1.0002.1111 Hypothetical Protein 4186568:4187446 forward
P1.0002.1112	0	-80,491353	6E-64	-42,323289	5,4E-44	0,02129668	9E-53	P1.0002.1112 aldehyde-activating protein 4187463:4187876 reverse
P1.0002.1113	0	-105,24604	0,0006	-53,682686	6,1E-19	0,01738459	4E-64	P1.0002.1113 aldehyde-activating protein 4187880:4188332 reverse
P1.0002.1114	0	-80,826324	4E-08	-42,263722	8,3E-45	0,02228239	2E-71	P1.0002.1114 hypothetical protein ABAZ39
P1.0002.1119	K01465	-3,9148401	3E-07	-2,1112935	0,02278	0,4368953	0,006	P1.0002.1119 dihydroorotase 4191651:4192955 forward
P1.0002.1123	K21700	-24,684679	0,0257	-13,040969	0,02113	0,09541415	0,004	P1.0002.1123 cupin 4196177:4196476 forward
P1.0002.1135	K00220	-40,390764	4E-08	-22,872705	0,0002	0,05402094	7E-06	P1.0002.1135 cyclohexadienyl dehydrogenase 4209197:4210069 reverse
P1.0002.1136	K00817	-9,5465037	0,0019	-4,867497	0,00032	0,19128414	2E-05	P1.0002.1136 histidinol-phosphate aminotransferase 4210255:4211349 reverse
P1.0002.1154	K06282	-6,3711926	0,0147	-3,2607185	0,00677	0,29347819	5E-09	P1.0002.1154 hydrogenase 2 small subunit 4226424:4227587 forward
P1.0002.1155	0	-6,3591589	3E-08	-3,3713639	1,2E-05	0,28413887	6E-09	P1.0002.1155 hydrogenase 4227595:4228650 forward
P1.0002.1156	0	-4,409481	0,0006	-2,3162092	0,00016	0,41523238	2E-06	P1.0002.1156 hydrogenase 4228663:4229817 forward
P1.0002.1157	K06281	-5,3997378	2E-08	-2,8515355	6,3E-07	0,33318691	4E-11	P1.0002.1157 hydrogenase 2 large subunit 4229836:4231539 forward
P1.0002.1161	K03192	-7,4561206	0,0039	-3,8506117	0,00133	0,25274595	6E-05	P1.0002.1161 urease accessory protein 4234595:4235197 reverse
P1.0002.1166	0	-7,8091194	0,002	-4,0710217	0,00011	0,24182689	4E-07	P1.0002.1166 Hypothetical Protein 4237781:4238122 reverse
P1.0002.1167	K04652	-7,3274459	0,0002	-3,8554027	8,6E-05	0,25523053	5E-07	P1.0002.1167 hydrogenase nickel incorporation protein HypB 4238177:4238857 forward
P1.0002.1169	K04656	-9,4241759	0,0004	-5,2577463	0,00275	0,19731788	3E-04	P1.0002.1169 carbamoyltransferase 4239462:4241729 forward
P1.0002.1178	K13581	-4,6321928	0,0004	-2,398883	0,00108	0,39049032	3E-05	P1.0002.1178 modification methylase 4248887:4249972 forward
P1.0002.1179	K01633	-5,3287519	4E-08	-2,9084788	0,00356	0,34913	2E-04	P1.0002.1179 diguanylate cyclase 4250106:4250489 forward
P1.0002.1185	K03529	-5,004258	0,0055	-2,5433544	0,03754	0,34440501	0,008	P1.0002.1185 chromosome segregation ATPase 4256435:4259896 forward
P1.0002.1208	K04562	-3,844066	4E-08	-2,0478747	0,04098	0,46553851	0,007	P1.0002.1208 cobyrinic acid a,c-diamide synthase 4288797:4289591 forward
P1.0002.1211	0	-5,262709	1E-05	-2,7761301	0,00026	0,34570024	9E-07	P1.0002.1211 hypothetical protein ABAZ39
P1.0002.1215	0	-17,046254	2E-10	-9,318536	0,0004	0,09775161	1E-10	P1.0002.1215 hypothetical protein 4296679:4296861 reverse
P1.0002.1234	0	-4,6144892	1E-05	-2,4374779	0,00016	0,39356226	1E-06	P1.0002.1234 C4-dicarboxylate ABC transporter 4315360:4316871 forward

P1.0002.1243	0	-3,907279	2E-20	-2,076814	0,00077	0,45294757	5E-05	P1.0002.1243 Hypothetical Protein 4328412:4328552 forward
P1.0002.1247	K09811	-10,140319	2E-06	-5,3075249	0,00037	0,17877989	1E-05	P1.0002.1247 membrane protein 4330865:4331767 forward
P1.0002.1274	0	-4,3647851	0,0002	-2,2445563	0,00751	0,40721089	7E-04	P1.0002.1274 hypothetical protein 4358153:4358923 reverse
P1.0002.1275	K08989	-13,785269	0,0108	-7,0395159	0,00166	0,14020345	6E-06	P1.0002.1275 membrane protein 4358970:4359395 reverse
P1.0002.1277	K01682	-4,4473239	2E-05	-2,3437915	8,5E-05	0,40743699	3E-07	P1.0002.1277 bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase 4361589:4364174 forward
P1.0002.1278	0	#DIV/0!	0,0009	#DIV/0!	0,03216	0,0191922	0,01	P1.0002.1278 AraC family transcriptional regulator 4364336:4365136 reverse
P1.0002.1279	K08223	-4,3544048	0,0071	-2,2366988	0,00656	0,42383684	1E-04	P1.0002.1279 Fosmidomycin resistance protein 4365264:4366469 forward
P1.0002.1283	K07220	-6,123139	7E-06	-3,2127843	0,00223	0,29578404	1E-04	P1.0002.1283 putative low-affinity phosphate transport accessory protein 4368816:4369469 forward
P1.0002.1284	K03306	-8,9331933	2E-08	-4,8459553	0,00033	0,21136808	4E-06	P1.0002.1284 inorganic phosphate transporter 4369474:4370481 forward
P1.0002.1322	K00311	-5,5047221	2E-12	-2,9411175	3,4E-06	0,32523599	4E-09	P1.0002.1322 electron transfer flavoprotein 4407884:4409533 reverse
P1.0002.1341	K01996	-10,255331	3E-05	-5,4791183	0,03288	0,17828707	0,007	P1.0002.1341 amino acid ABC transporter ATPase 4426056:4426760 reverse
P1.0002.1342	K07090	-14,908658	7E-08	-8,277041	5,9E-05	0,13523337	4E-07	P1.0002.1342 membrane protein 4426837:4427580 reverse
P1.0002.1343	K01895	-5,734944	1E-06	-3,1244679	0,00096	0,32416987	2E-05	P1.0002.1343 AMP-dependent synthetase 4427835:4429451 reverse
P1.0002.1375	K04750	-11,229604	5E-05	-5,9240529	0,00021	0,16880266	4E-06	P1.0002.1375 glyoxalase/bleomycin resistance protein/dioxygenase 4463638:4464072 reverse
P1.0002.1380	K00819	-55,511577	3E-09	-30,788007	1,1E-12	0,03727364	1E-18	P1.0002.1380 ornithine-oxoacid aminotransferase 4467546:4468763 reverse
P1.0002.1381	K01476	-24,694781	5E-15	-13,926892	2,2E-07	0,08097456	1E-10	P1.0002.1381 arginase 4468806:4469771 reverse
P1.0002.1390	K06287	-9,8886022	9E-05	-5,214274	0,00085	0,19175579	4E-05	P1.0002.1390 septum formation inhibitor Maf 4478331:4478936 reverse
P1.0002.1402	K02071	-8,0751811	3E-13	-4,3160867	6,8E-05	0,21724152	9E-07	P1.0002.1402 methionine ABC transporter ATP-binding protein 4489388:4490452 forward
P1.0002.1419	K01179	-5,3743898	0,0006	-2,7791403	0,00011	0,33712703	1E-06	P1.0002.1419 endoglucanase 4509212:4511698 forward
P1.0002.1446	0	-6,3212282	5E-06	-3,236838	0,02231	0,27025262	0,008	P1.0002.1446 Rrf2 family transcriptional regulator 4543766:4544218 reverse
P1.0002.1454	0	-3,7524773	7E-05	-2,002679	0,02642	0,48622537	0,002	P1.0002.1454 hypothetical protein 4550596:4553742 reverse
P1.0002.1472	K12950	-9,3877728	2E-06	-4,9144486	8,8E-13	0,19271366	3E-18	P1.0002.1472 ATPase 4568979:4571117 reverse
P1.0002.1475	0	-4,3107815	1E-35	-2,363358	0,00314	0,40334302	2E-04	P1.0002.1475 hypothetical protein ABAZ39

P1.0002.1496	K01251	-27,242417	2E-07	-14,624499	2,1E-17	0,06961909	4E-24	P1.0002.1496 adenosylhomocysteinase 4593040:4594344 forward
P1.0002.1501	K00821	-3,8203258	0,0006	-2,0014496	0,01184	0,47950036	5E-04	P1.0002.1501 acetylornithine aminotransferase 4597285:4598454 reverse
P1.0002.1502	K01992	-4,4319272	0,0003	-2,3471496	0,00708	0,41878032	3E-04	P1.0002.1502 multidrug ABC transporter permease 4598590:4599372 reverse
P1.0002.1505	K02902	-13,543804	2E-37	-7,2662846	1,5E-17	0,13128547	4E-25	P1.0002.1505 50S ribosomal protein L28 4601091:4601393 reverse
P1.0002.1507	K07588	-5,4853688	6E-08	-3,0111328	0,00968	0,32260109	9E-04	P1.0002.1507 G-protein chaperone required for conversion of methylmalonyl-CoA to succinyl-CoA 4602606:4603610 reverse
P1.0002.1509	K06147	-4,3255585	0,0002	-2,2271384	0,0111	0,40807061	0,001	P1.0002.1509 metal ABC transporter permease 4604288:4606147 reverse
P1.0002.1525	0	-4,8436795	4E-07	-2,6384299	0,01401	0,37100726	0,001	P1.0002.1525 hypothetical protein ABAZ39
P1.0002.1526	K08484	-5,1104073	1E-06	-2,8521323	0,02583	0,33687406	3E-04	P1.0002.1526 peptidase 4622921:4625209 reverse
P1.0002.1529	K00568	-13,910282	8E-06	-7,7480312	0,0142	0,14975448	0,001	P1.0002.1529 3-demethylubiquinone-9 3-methyltransferase 4627762:4628523 forward
P1.0002.1537	K03684	-7,1968698	1E-15	-3,996814	0,00018	0,25601721	3E-06	P1.0002.1537 3'-5' exonuclease 4635233:4635859 reverse
P1.0002.1538	K00329	-7,8387356	8E-10	-4,1973515	8,6E-05	0,23285361	4E-07	P1.0002.1538 3-beta hydroxysteroid dehydrogenase 4636299:4637288 forward
P1.0002.1557	0	-5,745461	5E-07	-3,0754767	0,01419	0,32440172	0,002	P1.0002.1557 hypothetical protein 4655583:4655789 forward
P1.0002.1559	K02051	-9,194409	1E-05	-4,8418406	5,2E-10	0,19960719	5E-14	P1.0002.1559 nitrate ABC transporter substrate-binding protein 4656597:4657574 reverse
P1.0002.1560	K02050	-4,1165035	8E-11	-2,226244	0,03018	0,43475507	0,005	P1.0002.1560 ABC transporter permease 4657582:4658328 reverse
P1.0002.1577	K00767	-4,9945793	2E-11	-2,7225999	0,0024	0,36534752	6E-05	P1.0002.1577 nicotinate-nucleotide pyrophosphorylase 4675895:4676728 forward
P1.0002.1589	0	-6,6037785	5E-16	-3,5531807	0,00691	0,27466269	6E-04	P1.0002.1589 hypothetical protein 4689030:4689239 forward
P1.0002.1597	0	-4,9201333	0,0001	-2,6173676	0,01929	0,37040223	0,002	P1.0002.1597 hypothetical protein ABAZ39
P1.0002.1636	0	-5,1455092	1E-06	-2,7139396	0,03631	0,34675474	0,008	P1.0002.1636 regulator 4744499:4745299 reverse
P1.0002.1639	0	-6,2092151	3E-08	-3,3366502	1,1E-06	0,29302023	6E-09	P1.0002.1639 chemotaxis protein 4748731:4750407 reverse
P2.0003.0009	0	-4,2575573	0,0002	-2,2246437	0,00054	0,42752074	1E-05	P2.0003.0009 SapC family protein 4779359:4780084 forward
P2.0003.0011	0	-10,956672	2E-09	-5,8779076	0,00758	0,17271961	8E-04	P2.0003.0011 hypothetical protein ABAZ39
P2.0003.0016	0	-4,7040732	6E-05	-2,5311135	0,02688	0,37847549	0,004	P2.0003.0016 hypothetical protein ABAZ39
P2.0003.0025	K02299	#DIV/0!	2E-14	#DIV/0!	6,2E-07	0,00459029	3E-09	P2.0003.0025 cytochrome o ubiquinol oxidase subunit III 4796778:4797413 forward
P2.0003.0026	K02300	-4,6162288	6E-06	-2,4516711	0,04857	0,39870455	0,008	P2.0003.0026 cytochrome C oxidase subunit III 4797410:4797781 forward
P2.0003.0030	0	-7,7873502	7E-08	-4,4146474	0,01907	0,25042058	0,003	P2.0003.0030 hypothetical protein 4800621:4800974 forward
P2.0003.0035	K02567	-7,5145552	0,0006	-3,9675465	0,00943	0,25167056	0,001	P2.0003.0035 nitrate reductase, partial 4803347:4805233 reverse

P2.0003.0112	K07395	-3,8119064	3E-05	-2,0031573	0,00164	0,47544901	4E-05	P2.0003.0112 peptidase 4885556:4886290 reverse
P2.0003.0170	0	-10,675785	2E-05	-5,8086543	0,00055	0,16797824	2E-05	P2.0003.0170 methanol dehydrogenase 4958495:4960378 reverse
P2.0003.0200	0	-5,4895926	1E-06	-2,9285708	0,00095	0,33568586	8E-06	P2.0003.0200 putative Tail Collar domain protein 4995160:4995846 reverse
P2.0003.0201	0	-5,0387773	4E-19	-2,7438819	0,0012	0,35843032	2E-05	P2.0003.0201 hypothetical protein 4995922:4996242 reverse
P2.0003.0202	0	-7,5775482	0,0005	-3,9272762	1,7E-05	0,24266052	6E-08	P2.0003.0202 putative Tail Collar domain protein 4996353:4997021 reverse
P2.0003.0203	0	-4,2687771	3E-06	-2,2113351	0,00903	0,41135388	9E-04	P2.0003.0203 acetyltransferase 4997433:4998002 forward
P2.0003.0208	K16922	-4,7305934	4E-07	-2,6447937	0,02896	0,36840445	4E-04	P2.0003.0208 peptidase 5015294:5017495 forward
P2.0003.0222	K03303	-5,0526149	0,0005	-2,6428756	2,3E-05	0,36310812	2E-07	P2.0003.0222 L-lactate permease 5032398:5034086 reverse
P2.0003.0223	K05799	-12,188501	9E-08	-6,5692283	0,00334	0,15449718	4E-04	P2.0003.0223 GntR family transcriptional regulator 5034473:5035243 forward
P2.0003.0231	0	-20,132857	0,0001	-10,735801	4,7E-07	0,09778583	1E-09	P2.0003.0231 ABC transporter substrate-binding protein 5042401:5043366 forward
P2.0003.0235	0	-17,354698	2E-05	-9,980211	0,01066	0,10281156	9E-07	P2.0003.0235 3-ketoacyl-ACP synthase 5046625:5047401 forward
P2.0003.0245	0	-6,1591554	8E-13	-3,3495222	0,0094	0,27453609	0,003	P2.0003.0245 hypothetical protein 5056161:5056562 forward
P2.0003.0249	0	-27,525783	0,0002	-14,117425	8,9E-13	0,06595016	3E-20	P2.0003.0249 nucleoside diphosphate kinase 5061368:5061760 reverse
P2.0003.0250	K00262	-48,767062	3E-12	-26,048933	9,1E-36	0,03748098	8E-50	P2.0003.0250 glutamate dehydrogenase 5062057:5063316 forward
P2.0003.0251	0	#DIV/0!	0,0002	#DIV/0!	1E-15	0,00127315	7E-25	P2.0003.0251 Hypothetical Protein 5063244:5063528 reverse
P2.0003.0252	0	-36,196749	0,0002	-19,001269	3,9E-09	0,05389321	2E-12	P2.0003.0252 two-component sensor histidine kinase 5063527:5065125 forward
P2.0003.0289	0	-6,6081447	5E-09	-3,6564557	0,00844	0,25891325	2E-04	P2.0003.0289 hypothetical protein ABAZ39
P2.0003.0293	0	-8,5966253	4E-05	-4,4170034	0,0066	0,20149946	0,001	P2.0003.0293 glycosyl transferase 5108139:5109440 reverse
P2.0003.0295	0	-5,8589947	0,0027	-3,0994942	0,03718	0,32571777	0,009	P2.0003.0295 glycosyl transferase family 2 5110269:5112620 reverse
P2.0003.0296	0	-13,986534	3E-05	-7,4048692	1,3E-07	0,13525962	6E-11	P2.0003.0296 hemolysin expression modulating protein 5112695:5113516 reverse
P2.0003.0335	K00253	-6,466519	1E-07	-3,4071993	7,9E-09	0,27897474	3E-14	P2.0003.0335 isovaleryl-CoA dehydrogenase 5168962:5170134 forward
P2.0003.0336	K00626	-9,4914327	1E-12	-5,0185537	1,5E-08	0,18669377	5E-13	P2.0003.0336 acetyl-CoA acetyltransferase 5170171:5171352 forward
P2.0003.0337	0	-7,5438527	4E-06	-3,9676197	7,7E-07	0,2412842	1E-10	P2.0003.0337 acyl-CoA dehydrogenase 5171471:5172604 forward
P2.0003.0338	K01969	-11,152312	7E-07	-5,8173234	1,3E-09	0,1603388	5E-14	P2.0003.0338 methylcrotonoyl-CoA carboxylase 5172633:5174240 forward
P2.0003.0339	K13766	-44,009028	4E-13	-24,596383	7,9E-08	0,04756629	1E-11	P2.0003.0339 enoyl-CoA hydratase 5174258:5175040 forward

P2.0003.0340	K01968	-5,8371022	3E-08	-3,1019078	0,00056	0,30598213	6E-06	P2.0003.0340 3-methylcrotonyl-CoA carboxylase 5175061:5177064 forward
P2.0003.0341	K01640	-19,764309	6E-10	-11,088595	7,8E-05	0,10284933	3E-07	P2.0003.0341 hydroxymethylglutaryl-CoA lyase 5177067:5177966 forward
P2.0003.0342	0	-8,1322805	3E-05	-4,2341039	0,00028	0,2227734	1E-05	P2.0003.0342 universal stress protein UspA 5177998:5178825 forward
P2.0003.0343	0	#DIV/0!	5E-05	#DIV/0!	0,00016	0,00774819	5E-06	P2.0003.0343 Hypothetical Protein 5178732:5180192 reverse
P2.0003.0344	0	-27,905404	0,0049	-14,966872	0,00014	0,07661905	3E-08	P2.0003.0344 Hypothetical Protein 5180205:5180522 forward
P2.0003.0345	0	-23,436509	2E-16	-12,156063	2,2E-07	0,06896894	3E-09	P2.0003.0345 3-hydroxy-2-methylbutyryl-CoA dehydrogenase 5180560:5181321 forward
P2.0003.0353	0	-4,08453	0,0015	-2,1238473	0,01079	0,44713795	7E-04	P2.0003.0353 Hypothetical Protein 5188632:5190704 forward
P2.0003.0354	K02471	-4,5539743	0,009	-2,3506323	0,03458	0,40902791	0,006	P2.0003.0354 ABC transporter 5190711:5192525 reverse
P2.0003.0372	0	-8,8113431	0,0003	-4,4838827	0,00122	0,20236081	1E-04	P2.0003.0372 Hypothetical Protein 5209410:5209640 reverse
P2.0003.0377	K02556	-3,9334291	0,0016	-2,0307689	0,00168	0,46457299	1E-04	P2.0003.0377 flagellar motor protein MotA 5212361:5213221 reverse
P2.0003.0378	K01596	-5,79693	0,0024	-2,9942867	9,9E-05	0,31797793	3E-11	P2.0003.0378 phosphoenolpyruvate carboxykinase 5213501:5215324 reverse
P2.0003.0379	K07226	-4,0359726	1E-09	-2,1329158	0,00334	0,43924031	1E-04	P2.0003.0379 heme iron utilization protein 5215660:5216451 reverse
P2.0003.0387	K08483	-5,055006	8E-10	-2,7126954	0,00014	0,35753083	5E-07	P2.0003.0387 phosphoenolpyruvate-protein phosphotransferase 5222278:5224050 forward
P2.0003.0406	0	-3,8940263	1E-09	-2,0995785	0,01679	0,459327	9E-04	P2.0003.0406 extracellular solute-binding protein, family 1 5248586:5249977 reverse
P2.0003.0407	0	-3,9140184	1E-06	-2,0682182	0,00059	0,46097834	3E-06	P2.0003.0407 exported protein of unknown function 5250126:5250869 reverse
P2.0003.0409	K07406	-4,8509446	9E-22	-2,6020735	6,5E-06	0,36639689	2E-09	P2.0003.0409 alpha gluctosidase galactosidase 5252596:5253891 forward
P2.0003.0414	0	-6,6709344	2E-22	-3,6201736	3,5E-05	0,25900219	3E-09	P2.0003.0414 histidine kinase 5257412:5259265 forward
P2.0003.0429	0	-7,1876547	3E-11	-3,959458	0,03836	0,25892348	0,008	P2.0003.0429 Hypothetical Protein 5271775:5272032 reverse
P2.0003.0455	0	-7,6300017	6E-05	-3,9852029	7,1E-05	0,23812891	5E-07	P2.0003.0455 putative Methyl-accepting chemotaxis protein 5305611:5307605 reverse
P2.0003.0456	0	-4,7271891	0,0027	-2,4508613	0,00038	0,39006864	6E-06	P2.0003.0456 Tl
P2.0003.0458	0	-16,188464	1E-11	-9,100321	0,00509	0,12745087	4E-04	P2.0003.0458 cytochrome c precursor 5310078:5310395 reverse
P2.0003.0474	0	-4,7892337	0,0036	-2,4427189	0,02405	0,37801552	0,006	P2.0003.0474 hypothetical protein ABAZ39
P2.0003.0477	0	#DIV/0!	0,0198	#DIV/0!	0,0049	0,01235851	4E-04	P2.0003.0477 signal peptide protein 5330314:5331003 forward
P2.0003.0480	0	-4,3588056	5E-08	-2,3047512	0,01629	0,40923154	0,002	P2.0003.0480 methyltransferase 5332470:5333063 forward

P2.0003.0486	K01999	-5,8009963	0,0002	-3,0308852	2,6E-07	0,31492229	1E-10	P2.0003.0486 amino acid ABC transporter substrate-binding protein 5340558:5341670 forward
P2.0003.0488	K01998	-15,14553	0,0013	-7,7599039	5,8E-07	0,12165331	8E-13	P2.0003.0488 branched-chain amino acid ABC transporter permease 5342626:5343654 forward
P2.0003.0489	K01995	-18,45682	8E-12	-10,04429	5,7E-05	0,10110279	1E-06	P2.0003.0489 ABC transporter 5343651:5344418 forward
P2.0003.0490	K01996	-5,2170487	5E-06	-2,7730053	0,04304	0,32391932	0,013	P2.0003.0490 leucine/isoleucine/valine transporter ATP-binding subunit 5344415:5345140 forward
P2.0003.0500	K00570	-9,2165416	0,0325	-4,5502678	0,03881	0,20414893	0,008	P2.0003.0500 phosphatidylethanolamine N-methyltransferase 5355369:5356004 reverse
P2.0003.0578	K01990	-4,3485815	6E-13	-2,3132381	0,00354	0,39599597	5E-04	P2.0003.0578 multidrug ABC transporter ATP-binding protein 5447976:5448902 reverse
P2.0003.0581	0	#DIV/0!	4E-05	#DIV/0!	0,03938	0,0203136	0,018	P2.0003.0581 hypothetical protein 5452058:5452252 reverse
P2.0003.0594	K00163	-5,1870234	0,0019	-2,6906672	0,00024	0,35505715	2E-06	P2.0003.0594 transketolase 5461034:5463442 forward
P2.0003.0600	K01665	-16,093183	0,0055	-7,8939092	0,02786	0,11491725	0,008	P2.0003.0600 aminobenzoate synthetase 5467180:5468541 reverse
P2.0003.0601	K00046	-4,8965405	5E-09	-2,641694	0,01927	0,36251855	0,003	P2.0003.0601 gluconate 5-dehydrogenase 5468571:5469356 reverse
P2.0003.0622	K00286	#DIV/0!	0,001	#DIV/0!	0,03097	0,01897603	0,007	P2.0003.0622 pyrroline-5-carboxylate reductase 5489762:5490628 reverse
P2.0003.0642	0	-8,0750661	0,0017	-4,1698597	0,00077	0,2306345	4E-05	P2.0003.0642 Zn-dependent protease with chaperone function 5513536:5514309 forward
P2.0003.0651	K08358	-5,5856189	2E-08	-2,9178419	0,01004	0,30783364	0,002	P2.0003.0651 tetrathionate reductase subunit B 5524438:5525193 forward
P2.0003.0654	K08357	-6,2585816	0,0005	-3,2713604	0,00043	0,29343482	6E-06	P2.0003.0654 tetrathionate reductase 5526340:5529447 forward
P2.0003.0666	K02000	-8,5178655	6E-37	-4,5836306	3,4E-12	0,20818578	3E-18	P2.0003.0666 glycine/betaine ABC transporter ATP-binding protein 5541884:5543080 forward
P2.0003.0667	K02001	-10,252317	1E-13	-5,4287225	2,2E-14	0,17430392	5E-21	P2.0003.0667 glycine betaine transporter subunit ; membrane component of ABC superfamily 5543083:5544021 forward
P2.0003.0668	K02002	-5,3801609	2E-112	-2,9167753	3,3E-06	0,32668337	1E-11	P2.0003.0668 glycine/betaine ABC transporter substrate-binding protein 5544063:5544974 forward
P2.0003.0686	0	-28,014891	7E-09	-15,802239	0,00361	0,07798137	2E-04	P2.0003.0686 Hypothetical Protein 5566280:5566765 forward
P3.0004.0028	K02824	-6,6079947	0,0031	-3,4722231	0,0177	0,29122669	0,003	P3.0004.0028 uracil transporter 5606778:5608088 reverse
P3.0004.0054	0	-12,36964	7E-18	-6,8934438	0,0006	0,1491114	4E-05	P3.0004.0054 putative response regulator receiver (CheY-like protein) 5634213:5634557 forward
P3.0004.0069	0	-30,500729	0,0123	-16,23884	0,00138	0,07223829	6E-07	P3.0004.0069 hypothetical protein 5654012:5654653 reverse

P3.0004.0070	0	-12,618585	4E-05	-6,5956308	3E-16	0,14510097	1E-22	P3.0004.0070 flagellar motor switch protein FliG 5654670:5655698 reverse
P3.0004.0071	0	-13,291617	2E-12	-7,6847127	0,00303	0,1467134	2E-04	P3.0004.0071 hypothetical protein ABAZ39
P3.0004.0072	K02419	-19,991468	0,0007	-10,315083	6,2E-10	0,09237037	5E-25	P3.0004.0072 flagellar biosynthetic protein flip 5656036:5656818 reverse
P3.0004.0073	0	-12,103329	0,0097	-6,0885335	0,00131	0,15330698	1E-06	P3.0004.0073 flagellar motor switch protein FliN 5656853:5657194 reverse
P3.0004.0074	K02411	-13,257117	3E-25	-7,0929945	4,3E-09	0,13374649	2E-13	P3.0004.0074 flagellar assembly protein FliH 5657182:5657913 reverse
P3.0004.0075	K02409	-9,912146	4E-08	-5,2445652	4,6E-14	0,18238823	2E-22	P3.0004.0075 flagellar M-ring protein FliF 5657913:5659613 reverse
P3.0004.0076	K02415	-12,050383	2E-06	-6,3237788	3,3E-14	0,15107605	1E-19	P3.0004.0076 flagellar basal body-associated protein FliL 5659852:5660301 forward
P3.0004.0077	0	-20,387063	3E-12	-10,803163	1,2E-24	0,08816244	9E-37	P3.0004.0077 flagellar motor switch protein FliM 5660319:5661425 forward
P3.0004.0078	0	-12,419052	3E-51	-6,7555932	1,9E-10	0,1444009	1E-15	P3.0004.0078 hypothetical protein 5661406:5662062 forward
P3.0004.0079	0	-26,69686	2E-06	-14,126445	8,6E-19	0,06978135	2E-25	P3.0004.0079 flagellar motor switch protein 5662066:5662788 forward
P3.0004.0080	0	-9,4475711	5E-11	-5,0677124	4,3E-06	0,19289754	5E-09	P3.0004.0080 putative Flagellar hook-length control protein 5662809:5664110 forward
P3.0004.0081	K02389	-26,080487	7E-05	-13,5204	3,4E-16	0,07025932	4E-24	P3.0004.0081 flagellar hook capping protein FlgD 5664126:5664776 forward
P3.0004.0093	0	-71,079597	6E-06	-35,407464	3E-14	0,02424007	3E-19	P3.0004.0093 hypothetical protein ABAZ39
P3.0004.0094	K02393	-30,773443	4E-60	-16,413573	1,3E-31	0,05723588	4E-47	P3.0004.0094 flagellar L-ring protein FlgH 5685164:5685907 reverse
P3.0004.0095	K02386	-20,046173	4E-05	-10,359371	9,7E-18	0,09031079	1E-24	P3.0004.0095 FlgA protein 5685945:5686676 reverse
P3.0004.0096	K02392	-39,064251	3E-07	-20,808669	3,2E-34	0,04716118	5E-46	P3.0004.0096 flagellar basal body rod protein FlgG 5686682:5687467 reverse
P3.0004.0097	K02391	-23,24223	2E-13	-12,318502	1,4E-26	0,07720681	3E-39	P3.0004.0097 flagellar basal body rod protein FlgG 5687486:5688214 reverse
P3.0004.0098	0	-13,172023	2E-20	-6,9712207	1,8E-12	0,13143979	9E-17	P3.0004.0098 hypothetical protein ABAZ39
P3.0004.0099	K02556	-119,3698	4E-05	-61,117249	3,7E-25	0,01518108	2E-41	P3.0004.0099 flagellar motor protein MotA 5690828:5691715 forward
P3.0004.0100	0	-49,945629	4E-07	-26,174572	7,3E-25	0,03669175	1E-32	P3.0004.0100 chemotaxis protein 5691725:5692675 forward
P3.0004.0101	0	#DIV/0!	0,0006	#DIV/0!	0,00032	0,00843214	3E-05	P3.0004.0101 transporter 5692682:5693890 reverse

P3.0004.0114	0	-4,3085081	2E-05	-2,3262376	0,03679	0,4238526	0,005	P3.0004.0114 Fis family transcriptional regulator 5701376:5703007 reverse
P3.0004.0122	0	-22,165307	0,0005	-12,054387	0,0006	0,09711315	1E-05	P3.0004.0122 Hypothetical Protein 5709278:5709823 forward
P3.0004.0160	0	-7,4617279	8E-05	-3,8710141	0,00227	0,24438127	2E-04	P3.0004.0160 hypothetical protein ABAZ39
P3.0004.0240	0	-4,9784024	1E-06	-2,627406	0,00052	0,36266255	3E-06	P3.0004.0240 ABC transporter substrate-binding protein 5837463:5838548 reverse
P3.0004.0263	0	-4,3810141	7E-12	-2,3354764	0,00405	0,40536966	1E-04	P3.0004.0263 cell cycle transcriptional regulator 5860309:5861103 forward
P3.0004.0270	0	-3,7454088	9E-13	-2,024994	0,00751	0,47845822	3E-04	P3.0004.0270 hypothetical protein 5864897:5865796 forward
P3.0004.0272	0	-10,577438	1E-05	-5,5864764	5,6E-06	0,17552571	8E-09	P3.0004.0272 MFS transporter 5865959:5867155 reverse
P3.0004.0274	0	-6,2219498	4E-14	-3,3900658	0,02588	0,26563124	3E-04	P3.0004.0274 conserved exported protein of unknown function 5868187:5868792 reverse
P3.0004.0275	0	-8,9721271	0,004	-4,5796101	0,00024	0,20608764	2E-07	P3.0004.0275 ATP synthase 5868789:5869250 reverse
P3.0004.0276	0	-4,3762869	4E-13	-2,3554161	0,04832	0,40199669	0,011	P3.0004.0276 flagellum-specific ATP synthase, partial 5869250:5869594 reverse
P3.0004.0277	0	-11,866297	2E-10	-6,4351414	5,3E-09	0,15808706	2E-13	P3.0004.0277 Hypothetical Protein 5869558:5870127 reverse
P3.0004.0278	K02387	-7,5422895	9E-07	-3,9698816	5,3E-10	0,24034234	8E-15	P3.0004.0278 flagellar basal body rod protein FlgG 5870253:5870645 forward
P3.0004.0279	K02388	-15,23752	4E-09	-8,0532169	3,2E-19	0,11872472	5E-28	P3.0004.0279 flagellar basal-body rod protein FlgC 5870666:5871085 forward
P3.0004.0280	K02408	-16,998924	0,0018	-8,7776676	1,3E-07	0,10941013	8E-24	P3.0004.0280 flagellar basal body protein FliE 5871100:5871465 forward
P3.0004.0281	K02420	-8,1382023	4E-185	-4,4095508	4,7E-07	0,21328002	7E-13	P3.0004.0281 flagellar biosynthesis protein FliQ 5871618:5871884 forward
P3.0004.0282	0	-22,213721	3E-14	-12,416447	4,8E-05	0,07858832	1E-16	P3.0004.0282 transglycosylase 5871886:5872656 forward
P3.0004.0285	K02401	-4,9035072	7E-07	-2,5840441	6,6E-06	0,36814146	3E-09	P3.0004.0285 flagellar biosynthesis protein 5875600:5876697 forward
P3.0004.0305	0	-9,8942626	1E-25	-5,4419374	0,00109	0,17234349	4E-07	P3.0004.0305 hypothetical protein 5895777:5896304 reverse
P3.0004.0306	0	-28,038322	3E-07	-15,466297	3,8E-08	0,07322432	4E-12	P3.0004.0306 flagellar rod assembly protein FlgJ 5896301:5896606 reverse
P3.0004.0307	K02394	-30,302759	7E-10	-15,87454	2,9E-30	0,05899006	3E-41	P3.0004.0307 flagellar basal body P-ring biosynthesis protein FlgA 5896631:5897767 reverse
P3.0004.0308	K02397	-20,482372	0,0015	-10,626177	1,6E-08	0,09092123	4E-28	P3.0004.0308 flagellar hook protein 5897808:5898872 reverse
P3.0004.0309	K02396	-16,379573	9E-27	-8,7873064	5,3E-20	0,10892434	2E-28	P3.0004.0309 flagellar hook protein 5898877:5900274 reverse
P3.0004.0310	K02390	-13,980489	1E-05	-7,29013	3,4E-18	0,13028062	6E-31	P3.0004.0310 flagellar hook protein 5900300:5901532 reverse
P3.0004.0313	0	-5,2815779	0,0047	-2,7079824	0,00563	0,34810151	8E-04	P3.0004.0313 hypothetical protein 5902983:5904026 reverse

P3.0004.0321	0	-10,879058	0,002	-5,4364975	0,00031	0,16191735	3E-05	P3.0004.0321 multidrug transporter 5911676:5912623 reverse
P3.0004.0322	K06212	-6,0320853	2E-10	-3,1660783	0,00078	0,28847437	5E-05	P3.0004.0322 formate transporter 5912756:5913673 reverse
P3.0004.0323	K04069	-5,8539747	8E-05	-3,0200497	0,00089	0,30428621	5E-05	P3.0004.0323 pyruvate formate lyase-activating protein 5913690:5914469 reverse
P3.0004.0327	0	#DIV/0!	0,0009	#DIV/0!	0,00427	0,01244505	2E-04	P3.0004.0327 Hypothetical Protein 5916395:5916538 forward
P3.0004.0348	0	-3,8256097	1E-10	-2,0850666	0,00914	0,46749962	5E-04	P3.0004.0348 hypothetical protein ABAZ39
P3.0004.0404	K06942	-7,8735413	3E-05	-4,1161802	7,1E-06	0,2312549	9E-09	P3.0004.0404 GTP-binding protein YchF 6004662:6005762 reverse
P3.0004.0408	K02335	-4,2143206	6E-08	-2,255907	0,00101	0,42745146	2E-05	P3.0004.0408 DNA polymerase 6009719:6012625 reverse
P3.0004.0410	0	-8,1654948	1E-06	-4,4398917	0,03907	0,22727173	0,011	P3.0004.0410 hypothetical protein ABAZ39
P3.0004.0424	K01754	-7,203343	9E-05	-3,9665001	0,02292	0,2616926	0,004	P3.0004.0424 threonine dehydratase 6028124:6029368 reverse
P3.0004.0434	K00640	-24,579123	5E-19	-13,198714	2,2E-09	0,07435193	2E-13	P3.0004.0434 serine O-acetyltransferase 6037114:6037998 reverse
P3.0004.0435	0	-11,47579	2E-08	-6,1510019	0,00218	0,14853065	5E-04	P3.0004.0435 glycosyl transferase 6038358:6039071 forward
P3.0004.0450	K01624	-4,5732764	1E-16	-2,4661668	0,00028	0,39075588	3E-06	P3.0004.0450 fructose-bisphosphate aldolase 6059368:6060447 reverse
P3.0004.0453	0	-4,9159621	7E-15	-2,6733193	0,00042	0,35755218	5E-06	P3.0004.0453 GntR family transcriptional regulator 6065067:6066515 forward
P3.0004.0454	0	-15,807674	4E-07	-8,2890238	8,3E-15	0,1144698	1E-20	P3.0004.0454 nitrogen fixation protein FixG 6066705:6068201 forward
P3.0004.0455	0	-25,082324	0,0009	-12,49484	4,5E-08	0,06955996	8E-14	P3.0004.0455 membrane protein 6068208:6069194 reverse
P3.0004.0463	0	-3,8318944	5E-07	-2,0246781	0,01479	0,46895241	7E-04	P3.0004.0463 phenylacetic acid degradation protein PaaI 6075289:6075732 reverse
P3.0004.0464	K03149	-5,2408379	5E-06	-2,725396	0,01541	0,33579108	0,002	P3.0004.0464 thiazole synthase 6075783:6076553 reverse
P3.0004.0478	0	-5,2711768	9E-16	-2,8491888	0,02515	0,32873605	0,006	P3.0004.0478 hypothetical protein 6090518:6090775 reverse
P3.0004.0498	K01006	-6,5404363	3E-07	-3,4602986	8,3E-09	0,27627207	6E-14	P3.0004.0498 pyruvate phosphate dikinase 6109343:6112036 forward
P3.0004.0505	0	-13,360196	0,003	-6,8660522	0,04782	0,13021316	0,017	P3.0004.0505 Aby 6118464:6119594 forward
P3.0004.0508	K00948	-5,1360223	7E-07	-2,7112853	3,6E-05	0,35189645	1E-07	P3.0004.0508 ribose-phosphate pyrophosphokinase 6121205:6122140 forward
P3.0004.0509	K02897	-69,18709	4E-69	-38,203977	2,7E-19	0,0271542	2E-27	P3.0004.0509 50S ribosomal protein L25 6122451:6123062 forward
P3.0004.0535	K18306	-4,9259898	0,0015	-2,5813425	0,01244	0,37915641	0,001	P3.0004.0535 RND transporter 6151693:6152889 forward
P3.0004.0536	K18307	-4,7360511	0,0007	-2,4720648	9,5E-05	0,38673851	1E-06	P3.0004.0536 acriflavine resistance protein B 6152900:6155977 forward
P3.0004.0552	0	-12,198313	3E-08	-7,0392802	0,02283	0,16313274	0,003	P3.0004.0552 cupin 6175789:6176124 forward
P4.0005.0079	0	#DIV/0!	1E-06	#DIV/0!	0,00549	0,01302405	4E-04	P4.0005.0079 exported protein of unknown function 6300209:6300751 reverse
P5.0006.0010	0	-6,1271328	2E-07	-3,2730923	0,00599	0,29542723	3E-04	P5.0006.0010 Hypothetical Protein 6438290:6439459 reverse

P5.0006.0021	0	-6,8759556	2E-13	-3,6839958	2,4E-09	0,26075106	4E-14	P5.0006.0021 glycosyl transferase 6446468:6447952 reverse
P5.0006.0024	0	-8,78104	0,0009	-5,0108455	0,01694	0,22219858	0,002	P5.0006.0024 Hypothetical Protein 6450605:6452278 reverse
P5.0006.0026	0	-16,48275	2E-05	-8,6588781	4,4E-13	0,1118975	5E-19	P5.0006.0026 Hypothetical Protein 6453639:6457538 forward
P5.0006.0027	0	-5,0200125	9E-06	-2,8077914	0,03035	0,34840721	0,008	P5.0006.0027 Hypothetical Protein 6457535:6458620 forward
P5.0006.0028	0	-4,8459956	9E-32	-2,6370827	0,00093	0,36124008	2E-05	P5.0006.0028 Hypothetical Protein 6458617:6459078 reverse
P5.0006.0030	0	-5,4793806	0,0006	-2,8423809	1,7E-06	0,33429404	7E-10	P5.0006.0030 Hypothetical Protein 6460428:6461759 forward
P5.0006.0031	K09691	-3,9610612	0,0001	-2,0638527	0,00109	0,45717211	7E-05	P5.0006.0031 putative ABC-2 transporter ATP-binding component-like protein 6462110:6463171 forward
P5.0006.0034	0	-5,2757591	1E-22	-2,8732071	0,00085	0,33047363	3E-05	P5.0006.0034 glycosyl transferase-like protein 6465221:6465805 reverse
P5.0006.0035	0	-5,5180785	0	-2,9983146	1,7E-06	0,31740616	7E-12	P5.0006.0035 putative outer membrane secretion protein-like protein 6466988:6468016 forward
P5.0006.0038	0	-4,8789186	6E-06	-2,5417146	0,00317	0,36247719	3E-04	P5.0006.0038 glycosyl transferase family 1 6473127:6474374 forward
P5.0006.0043	0	-6,5533317	0,0002	-3,5461001	0,00217	0,29265356	1E-04	P5.0006.0043 glycosyl transferase-like protein 6480007:6481689 forward
P5.0006.0045	K01710	-5,7126166	0,0006	-2,9664554	8E-05	0,320916	5E-07	P5.0006.0045 F349575
P5.0006.0055	K01711	-6,0234991	0,0004	-3,1355314	1,6E-07	0,30395933	4E-12	P5.0006.0055 GDP-mannose dehydratase 6495990:6497051 reverse
P5.0006.0060	0	-4,7602297	2E-05	-2,4646468	3,3E-05	0,3823305	4E-06	P5.0006.0060 Hypothetical Protein 6504146:6505120 forward
P5.0006.0061	0	-4,9284445	5E-05	-2,5708866	5,7E-06	0,3689413	6E-09	P5.0006.0061 MULTISPECIES: hypothetical protein 6505176:6506195 forward
P5.0006.0065	0	-3,7346471	7E-16	-2,0196468	0,0016	0,47542269	5E-05	P5.0006.0065 hypothetical protein pRhico013 6508128:6510203 forward
P5.0006.0067	0	-3,9962234	0,0075	-2,0334406	0,03586	0,4562139	0,002	P5.0006.0067 hypothetical protein pRhico015 6512790:6513872 reverse
P5.0006.0069	0	-5,2318985	2E-06	-2,789071	0,0256	0,34208866	0,004	P5.0006.0069 glycosyl transferase-like protein 6515570:6516595 forward
P5.0006.0070	0	-6,5173001	6E-09	-3,5175762	2,3E-05	0,2711253	5E-08	P5.0006.0070 Hypothetical Protein 6516630:6519473 forward
P5.0006.0074	0	-5,3266378	0,0174	-2,7239895	0,01479	0,35094554	5E-07	P5.0006.0074 hypothetical protein pRhico022 6522709:6524538 reverse
P5.0006.0083	0	-5,5241211	2E-27	-2,9831316	5,3E-07	0,32155142	2E-10	P5.0006.0083 plasmid replication protein 6532870:6533808 forward
P5.0006.0087	K11005	-12,479119	1E-05	-6,9162218	0,00315	0,13641647	1E-06	P5.0006.0087 Hypothetical Protein 6537756:6540578 reverse
P5.0006.0094	K00012	-5,980365	0,0097	-3,0601365	0,00341	0,31124461	1E-07	P5.0006.0094 putative UDP glucose dehydrogenase 6547035:6548360 reverse

P5.0006.0095	0	-10,780858	3E-34	-5,8084987	1,5E-09	0,16594981	1E-14	P5.0006.0095 hypothetical protein ABAZ39
P5.0006.0097	K15894	-4,2522349	5E-08	-2,2500692	0,00026	0,4225236	3E-06	P5.0006.0097 putative sugar nucleotide epimerase/dehydratase 6550257:6551309 reverse
P5.0006.0107	0	-7,5962109	2E-16	-4,0961898	1E-08	0,23385037	2E-12	P5.0006.0107 secreted sugar hydrolase-like protein (modular protein) 6560975:6564319 reverse
P5.0006.0112	0	-28,009331	9E-10	-14,135026	0,00076	0,06396455	5E-05	P5.0006.0112 Hypothetical Protein 6569430:6569585 forward
P6.0001.0001	0	-5,0334409	7E-11	-2,7981676	0,00866	0,35216264	7E-04	P6.0001.0001 integrase 24:911 reverse
P6.0001.0077	0	-34,122516	2E-15	-17,477598	0,00011	0,05161768	9E-06	P6.0001.0077 Hypothetical Protein 39081:39461 reverse
P6.0001.0080	0	-7,5517084	3E-09	-4,0575025	0,01968	0,23405577	0,005	P6.0001.0080 hypothetical protein APCd
P6.0001.0082	0	-4,7014192	2E-20	-2,4921316	7,3E-06	0,37905789	8E-08	P6.0001.0082 hypothetical protein APCd
P6.0001.0087	0	-4,4838724	2E-16	-2,4916519	0,02463	0,38592018	7E-04	P6.0001.0087 hypothetical protein APCd
P6.0002.0007	0	-6,4426331	0,0035	-3,3463599	0,00426	0,29339783	4E-04	P6.0002.0007 hypothetical protein APCd
P6.0003.0036	0	-3,9813819	0,0019	-2,0708695	0,03959	0,4624463	0,006	P6.0003.0036 Hypothetical Protein 92490:93173 reverse
P6.0003.0042	0	#DIV/0!	0,0052	#DIV/0!	0,0465	0,02123384	0,012	P6.0003.0042 agmatinase 96867:97937 reverse
P6.0013.0004	K07218	-24,028217	0,0013	-12,305581	1,1E-07	0,07744801	5E-13	P6.0013.0004 carbohydrate-binding protein 237741:239108 reverse
P6.0013.0005	K00376	-9,2967261	0,0005	-4,8134723	1,6E-08	0,19744766	1E-16	P6.0013.0005 nitrous-oxide reductase 239249:240916 reverse
P6.0019.0008	0	#DIV/0!	9E-05	#DIV/0!	0,0389	0,02019716	0,015	P6.0019.0008 hypothetical protein 319539:320051 forward
P6.0019.0009	0	-4,2471497	0,0006	-2,2195707	0,01687	0,43234766	0,001	P6.0019.0009 Hypothetical Protein 320093:320761 reverse
P6.0022.0004	K03325	#DIV/0!	0,001	#DIV/0!	0,0139	0,01568097	0,002	P6.0022.0004 arsenic transporter 337031:338104 reverse
P6.0022.0015	0	-8,0468132	0,0022	-4,2862393	0,02515	0,21276654	0,009	P6.0022.0015 acyl-CoA synthetase 345071:346696 forward
P6.0025.0001	0	-15,873899	0,0043	-7,7602887	0,00148	0,11110099	4E-04	P6.0025.0001 Transposase, TnpA family 364382:364795 forward
P6.0025.0002	0	-5,8801294	2E-67	-3,1330874	5,9E-07	0,30209323	3E-10	P6.0025.0002 Hypothetical Protein 365296:365487 forward
P6.0030.0006	0	-5,0311536	8E-10	-2,7965695	0,02145	0,33975747	8E-06	P6.0030.0006 Hypothetical Protein 414335:414889 forward
P6.0035.0001	0	-4,37216	5E-06	-2,3674267	0,03196	0,41625794	0,003	P6.0035.0001 histidine kinase 464661:466118 forward
P6.0036.0002	K03704	-7,9741424	2E-05	-4,4169857	0,03759	0,20782872	0,002	P6.0036.0002 cold-shock protein 476797:477006 reverse
P6.0036.0005	0	-20,389708	2E-11	-11,742529	0,01898	0,10400108	0,004	P6.0036.0005 Hypothetical Protein 478578:478775 forward
P6.0036.0007	0	-98,895201	9E-06	-51,725024	3,4E-42	0,01857895	2E-69	P6.0036.0007 fimbrial protein 479817:480005 forward
P6.0036.0008	0	-84,821653	2E-146	-45,819786	7,9E-20	0,02072293	1E-36	P6.0036.0008 pilus assembly protein CpaA 480116:480664 forward
P6.0036.0009	K02279	-52,937162	0,0002	-27,825289	3,2E-14	0,03640968	1E-27	P6.0036.0009 pilus assembly protein CpaB 480668:481522 forward
P6.0036.0010	K02280	-42,680236	3E-05	-23,072412	1,2E-19	0,0462053	7E-28	P6.0036.0010 fimbrial protein 481624:483000 forward
P6.0036.0011	K02281	-8,1186204	6E-07	-4,234171	0,00013	0,21954484	2E-06	P6.0036.0011 hypothetical protein ABAZ39
P6.0036.0012	K02282	-15,621049	2E-05	-8,3688651	2,2E-08	0,123445	5E-12	P6.0036.0012 fimbrial protein 483688:484887 forward
P6.0036.0013	K02283	-107,31209	1E-07	-56,040029	2,6E-18	0,01768211	4E-26	P6.0036.0013 Hypothetical Protein 484892:485950 forward
P6.0048.0002	K08303	-11,350002	0,0021	-5,874632	0,00055	0,16778423	3E-05	P6.0048.0002 protease 533352:534293 reverse

P6.0061.0006	K04101	#DIV/0!	0,0002	#DIV/0!	0,00636	0,01338992	6E-04	P6.0061.0006 protocatechuate 3,4-dioxygenase subunit beta 588097:588939 forward
P6.0062.0001	K12511	-29,241043	0,0031	-15,745549	6E-05	0,07451813	1E-06	P6.0062.0001 fimbrial protein 593528:594490 forward
P6.0062.0002	0	-70,04735	0,0002	-34,650968	1,8E-08	0,02627521	4E-11	P6.0062.0002 hypothetical protein ABAZ39
P6.0062.0003	0	-24,679539	0,0041	-12,819675	1E-05	0,07907459	1E-13	P6.0062.0003 hypothetical protein ABAZ39
P6.0062.0004	0	-69,91986	0,0029	-37,580663	5,7E-07	0,03127214	6E-15	P6.0062.0004 hypothetical protein ABAZ39
P6.0062.0005	0	-14,351838	7E-06	-7,5539449	3,3E-12	0,12836656	7E-18	P6.0062.0005 hypothetical protein ABAZ39
P6.0062.0007	0	-14,649134	0,0003	-7,2689736	0,03943	0,12426714	0,011	P6.0062.0007 hypothetical protein ABAZ39
P6.0062.0008	K02575	-104,74479	5E-15	-58,478309	1,4E-21	0,01962779	6E-31	P6.0062.0008 nitrate/nitrite transporter NarK 599424:600806 reverse
P6.0062.0009	0	-14,494023	0,0109	-7,0469593	0,046	0,1289179	0,021	P6.0062.0009 hypothetical protein ABAZ39
P6.0093.0001	0	-5,7369993	3E-08	-3,1690687	0,01606	0,30608088	0,003	P6.0093.0001 transposase, partial 646637:647362 forward
P6.0105.0001	K04748	-17,481683	1E-09	-10,143594	0,03709	0,11988714	0,005	P6.0105.0001 BBQ
P6.0141.0001	0	-12,372517	8E-06	-6,3829706	0,0001	0,14512878	4E-06	P6.0141.0001 secretion system protein E 666169:666582 forward
P6.0368.0001	0	-4,7935333	0,0034	-2,4895241	0,00102	0,38421989	2E-07	P6.0368.0001 Hypothetical Protein 709195:709449 forward
P6.0663.0001	0	-5,8219847	0,0017	-2,9565666	0,00963	0,30591645	0,003	P6.0663.0001 Hypothetical Protein 745467:746282 reverse
P6.1664.0001	0	-4,5141571	6E-09	-2,355434	0,00478	0,38542873	8E-04	P6.1664.0001 Hypothetical Protein 863457:863714 reverse

Anexo 02: lista dos genes ativados de *A. brasilense* FP2 em condição de Fixação Biológica de Nitrôgenio.

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
c1.0001.0172	K09826	2,505543391	8,7E-06	4,736882712	6,6E-06	4,462084325	2E-13	c1.0001.0172 Fur family transcriptional regulator 166270:166701 forward
c1.0001.0261	K03089	3,732902613	1,6E-12	6,911349176	2,2E-14	6,499210746	4E-21	c1.0001.0261 RNA polymerase sigma 70 262019:262921 reverse
c1.0001.0278	K03320	58,25477556	0	109,1478938	2E-54	103,3472282	8E-109	c1.0001.0278 ammonia channel precursor 281406:282722 forward
c1.0001.0315	0	2,810844168	1,4E-09	5,385195837	0,00023	5,126241171	1E-23	c1.0001.0315 Hypothetical Protein 310509:310658 forward
c1.0001.0344	K11754	2,556220078	0,00014	4,878622278	2,3E-10	4,612054588	6E-15	c1.0001.0344 folylpolyglutamate synthase 348089:349387 forward
c1.0001.0544	K04566	2,57526338	0,00027	5,009861802	4,1E-13	4,771592933	6E-18	c1.0001.0544 lysyl-tRNA synthetase 573757:575325 reverse
c1.0001.0621	0	#DIV/0!	0,01381	Inf	0,00944	71,11726646	0,0019	c1.0001.0621 Hypothetical Protein 642149:642262 reverse
c1.0001.0636	0	#DIV/0!	0,0169	Inf	0,00857	39,2024627	0,0242	c1.0001.0636 conserved exported protein of unknown function 655244:655909 forward
c1.0001.0640	0	6,67139688	0,00029	13,6360077	6,4E-06	12,67443624	3E-06	c1.0001.0640 hydrolase 659952:660434 forward
c1.0001.0641	0	6,855356891	0	12,94319927	3,3E-17	12,09997491	6E-23	c1.0001.0641 hypothetical protein 660531:660779 reverse
c1.0001.0766	K09786	2,708448297	0,00103	5,113811032	0,00022	4,871620908	1E-22	c1.0001.0766 hypothetical protein ABAZ39
c1.0001.0775	0	2,029159997	0,0006	3,647840262	2E-09	3,458554733	5E-10	c1.0001.0775 hypothetical protein 802940:803152 reverse
c1.0001.0782	0	8,086902615	0,02922	15,18167029	0,0007	12,87866461	0,0003	c1.0001.0782 membrane protein 808573:809685 reverse
c1.0001.0795	0	2,341658456	0,0016	4,487836759	3,7E-12	4,285285766	2E-15	c1.0001.0795 two-component sensor histidine kinase 824840:827035 forward
c1.0001.0796	0	2,520185105	7,9E-13	4,719003192	4,9E-07	4,417488591	2E-09	c1.0001.0796 Hypothetical Protein 827165:827254 forward
c1.0001.0902	0	16,90653085	0	31,84279318	5E-29	30,29802679	3E-79	c1.0001.0902 hypothetical protein 940917:941402 reverse
c1.0001.0903	0	3,737329157	2E-14	6,990858172	5,3E-18	6,646623479	2E-30	c1.0001.0903 hypothetical protein 941672:942208 reverse
c1.0001.0940	0	5,611645197	0,00098	10,26952895	5,1E-11	9,746445967	9E-24	c1.0001.0940 chemotaxis protein 975841:976947 forward
c1.0001.0959	K03855	27,70668232	0	51,25406285	5,4E-53	48,25414072	8E-78	c1.0001.0959 FixX 1004051:1004344 reverse
c1.0001.0960	K00313	88,67840813	0	164,8397153	8,5E-52	155,6563419	1E-121	c1.0001.0960 protein fixC 1004341:1005642 reverse
c1.0001.0961	K03522	126,6588572	0	235,6108635	4,7E-73	219,4578233	3E-112	c1.0001.0961 electron transfer flavoprotein subunit alpha 1005673:1006752 reverse
c1.0001.0962	K03521	104,6927314	0	199,2518594	1,9E-69	188,2832639	1E-111	c1.0001.0962 FixA 1006806:1007657 reverse
c1.0001.0963	K02595	115,0988562	3,5E-10	223,695343	2E-73	212,410998	2E-103	c1.0001.0963 nitrogen fixation protein NifW 1008000:1008323 reverse
c1.0001.0964	K00640	57,73352793	0	107,4541061	4E-46	96,71889204	3E-53	c1.0001.0964 serine acetyltransferase 1008341:1009225 reverse
c1.0001.0965	K02594	33,68103215	0	61,36798472	3,9E-51	56,66606888	4E-63	c1.0001.0965 homocitrate synthase 1009239:1010384 reverse

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
c1.0001.0966	K04487	206,4274595	0	373,8281176	3,2E-73	344,2014375	6E-125	c1.0001.0966 cysteine desulfurase 1010653:1011858 reverse
c1.0001.0967	K13819	190,0864722	0	320,6451905	9,1E-76	287,6907389	3E-109	c1.0001.0967 nitrogen fixation protein NifU 1011862:1012779 reverse
c1.0001.0968	K13628	257,2632785	0	439,4163283	3,8E-74	374,2833809	6E-86	c1.0001.0968 hypothetical protein 1012828:1013208 reverse
c1.0001.0970	0	8,204995711	1,1E-06	16,41263621	7,4E-19	15,8359094	1E-22	c1.0001.0970 hydrogenase 1014022:1014612 reverse
c1.0001.0971	0	25,34988318	1,9E-08	50,34298439	1E-40	48,05103848	3E-47	c1.0001.0971 ferredoxin 1014665:1014958 reverse
c1.0001.0972	0	276,841529	0	432,1178973	2E-73	356,9975133	3E-83	c1.0001.0972 F361867
c1.0001.0973	0	85,372915	0	145,4531791	5E-66	130,519328	3E-81	c1.0001.0973 hypothetical protein 1015348:1015812 reverse
c1.0001.0974	K02596	152,1081107	0	335,3407157	6,4E-67	315,4591386	2E-72	c1.0001.0974 nitrogen fixation protein NifX 1015817:1016287 reverse
c1.0001.0975	K02592	202,0251296	4,1E-11	365,9601644	1,6E-78	332,896663	3E-105	c1.0001.0975 NifN
c1.0001.0976	K02587	93,17852973	0	170,2873528	3,4E-69	157,9597026	1E-102	c1.0001.0976 nitrogenase iron-molybdenum cofactor biosynthesis protein NifE 1017672:1019087 reverse
c1.0001.0977	K03406	2,455731521	0,00079	4,774085945	5,8E-13	4,545415713	2E-16	c1.0001.0977 chemotaxis protein 1019277:1020626 reverse
c1.0001.0979	0	9,426464053	8,2E-10	17,50823995	1,3E-10	15,87264461	3E-13	c1.0001.0979 transcriptional regulator, BadM/Rrf2 family (fragment), partial 1022559:1023089 reverse
c1.0001.0980	K02591	242,6347503	0	442,3659963	4E-30	417,3718783	1E-150	c1.0001.0980 nitrogenase molybdenum-iron protein beta chain 1023301:1024860 reverse
c1.0001.0981	K02586	324,880163	0	582,2220517	7,6E-31	546,342929	3E-141	c1.0001.0981 nitrogenase molybdenum-iron protein alpha chain 1024891:1026330 reverse
c1.0001.0982	K02588	159,4925342	0	285,8647443	1,6E-50	266,9604012	1E-121	c1.0001.0982 nitrogenase iron protein, nifH; dinitrogenase reductase 1026471:1027352 reverse
c1.0001.0988	K01118	5,112800242	2,4E-09	9,392716239	1,1E-07	8,675654225	8E-18	c1.0001.0988 azoreductase 1032677:1033303 forward
c1.0001.0989	0	12,47182051	0	22,75974888	2,8E-38	21,50884434	3E-56	c1.0001.0989 F216815
c1.0001.0990	0	18,48937464	0	36,01892044	3,7E-48	34,45130367	7E-62	c1.0001.0990 UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferase 1034414:1035595 reverse
c1.0001.0991	0	30,46852906	0	51,13891665	5,7E-37	45,23093374	8E-49	c1.0001.0991 hypothetical protein 1035610:1035894 reverse
c1.0001.0992	K02593	9,801735643	0	18,86250166	4,8E-21	17,81946447	4E-37	c1.0001.0992 protein fixU 1035900:1036109 reverse
c1.0001.0993	K04487	15,087303	0	28,33504154	7,1E-43	26,82421776	3E-57	c1.0001.0993 cysteine desulfurase 1036151:1037302 reverse
c1.0001.0994	K02597	16,64671944	0,00013	30,72386676	2,3E-16	28,9535819	2E-40	c1.0001.0994 nitrogen fixation protein NifZ 1037337:1037648 reverse
c1.0001.0995	K02597	18,59997364	0	37,23609306	2,3E-46	35,68997757	7E-54	c1.0001.0995 nifZ, nitrogenase P-cluster assembly 1037672:1037968 reverse
c1.0001.0996	0	24,64899304	1,2E-10	47,92423851	1,4E-49	45,71904262	2E-58	c1.0001.0996 hypothetical protein ABAZ39

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
c1.0001.0997	0	22,08587443	0	40,95234639	7,4E-48	38,45710886	2E-61	c1.0001.0997 hypothetical protein 1038763:1039296 reverse
c1.0001.0998	0	8,957993541	6,2E-05	16,80416404	7,3E-19	16,01554012	4E-43	c1.0001.0998 ferredoxin 1039341:1039532 reverse
c1.0001.0999	K02585	72,61311381	0	131,9448075	4,6E-59	123,6121564	2E-105	c1.0001.0999 FeMo cofactor biosynthesis protein NifB 1039612:1041171 reverse
c1.0001.1003	K02584	4,06593665	0,00106	7,641344156	5E-13	7,306433228	8E-17	c1.0001.1003 Hypothetical Protein 1044715:1045485 forward
c1.0001.1059	K15790	9,933506848	0,00013	18,84539022	2E-16	17,99724526	5E-38	c1.0001.1059 hydrogenase 1107819:1108445 reverse
c1.0001.1213	K07234	2,045333739	0,00274	3,815748595	9,6E-08	3,619980393	2E-11	c1.0001.1213 short-chain dehydrogenase 1268751:1269959 reverse
c1.0001.1253	K07576	2,791999651	2,4E-10	5,307224338	2,1E-14	5,044887117	1E-20	c1.0001.1253 hypothetical protein 1302469:1303212 forward
c1.0001.1254	K07576	3,367424241	0,00019	6,133319841	6,1E-17	5,826677863	2E-20	c1.0001.1254 beta-lactamase 1303176:1304075 forward
c1.0001.1384	0	5,769886568	0	10,7111643	2,3E-22	10,11389434	1E-29	c1.0001.1384 hypothetical protein ABAZ39
c1.0001.1464	K00788	6,817946202	0,00033	13,20438874	4,6E-13	12,60750376	9E-15	c1.0001.1464 thiamine-phosphate pyrophosphorylase 1530046:1530666 forward
c1.0001.1474	0	2,966089633	9,4E-06	5,515223215	1,4E-11	5,205613795	3E-17	c1.0001.1474 Fis family transcriptional regulator 1538567:1540108 reverse
c1.0001.1475	K00285	2,584327932	0,00087	4,853460727	2,1E-13	4,62571052	2E-19	c1.0001.1475 D-amino acid dehydrogenase small subunit 1540381:1541694 forward
c1.0001.1476	0	4,127311941	9,6E-06	7,661153214	2,4E-09	7,234811059	1E-11	c1.0001.1476 hypothetical protein ABAZ39
c1.0001.1482	K03190	6,276203871	0,00175	11,08491536	0,00107	9,614319708	1E-06	c1.0001.1482 Urease accessory protein UreD 1548838:1549665 forward
c1.0001.1495	0	3,403934477	8,4E-07	6,523131447	1,5E-17	6,219442276	8E-26	c1.0001.1495 N-formylglutamate amidohydrolase 1558391:1559215 reverse
c1.0001.1525	0	2,066082144	0,01838	3,96992122	5,3E-05	3,793274169	4E-06	c1.0001.1525 hypothetical protein 1589310:1589561 reverse
c1.0001.1603	K09970	2,029807892	0,00537	3,857079607	2,7E-10	3,677367159	7E-14	c1.0001.1603 ABC transporter permease 1661260:1662459 forward
c1.0001.1615	0	3,713447278	0,00911	7,426323906	0,01519	6,679923329	0,0112	c1.0001.1615 hypothetical protein 1672263:1672439 forward
c1.0001.1684	0	2,23569738	0,01227	4,290655241	1,3E-05	4,108019798	4E-07	c1.0001.1684 Hypothetical Protein 1748138:1748362 forward
c1.0001.1763	0	3,091979091	2,4E-06	5,76766232	1,1E-15	5,486326949	3E-21	c1.0001.1763 Hypothetical Protein 1824938:1825237 forward
c1.0001.1892	0	5,743601933	3,2E-11	10,51320141	3E-24	9,978920779	9E-37	c1.0001.1892 Nex18 symbiotically induced protein 1960673:1961161 reverse
c1.0001.1901	0	3,509122693	0,00237	6,076711506	2,6E-06	5,6212608	4E-07	c1.0001.1901 hypothetical protein 1972104:1972337 reverse
c1.0001.1935	0	2,824882015	0,00851	5,001619449	0,00302	4,654388195	1E-08	c1.0001.1935 Hypothetical Protein 2004704:2004964 reverse
c1.0001.1989	0	2,324071324	0,01376	4,463754397	0,01832	4,048010693	0,0105	c1.0001.1989 hypothetical protein 2056001:2056135 reverse
c1.0001.2098	K02408	2,150135844	0,00583	3,939390075	1,9E-10	3,747117288	1E-12	c1.0001.2098 flagellar hook-basal body protein FliE

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
								2177250:2177549 reverse
c1.0001.2203	K11963	14,71629139	0	27,78273343	2,6E-32	25,92721691	4E-42	c1.0001.2203 urea ABC transporter ATP-binding protein 2274885:2275580 reverse
c1.0001.2204	K11962	23,37554793	0	42,88650062	2,6E-25	38,30090436	7E-32	c1.0001.2204 urea ABC transporter ATP-binding protein 2275741:2276487 reverse
c1.0001.2205	K11961	42,09613469	0	88,01955782	4E-35	81,89428003	9E-42	c1.0001.2205 amino acid ABC transporter permease 2276493:2277632 reverse
c1.0001.2206	K11960	16,37341267	8,1E-09	28,1516618	5E-19	24,9895332	2E-23	c1.0001.2206 urea ABC transporter permease 2277637:2279241 reverse
c1.0001.2207	K11959	31,47741171	0	59,62331456	1,5E-55	56,30712015	1E-76	c1.0001.2207 branched-chain amino acid ABC transporter substrate-binding protein 2279415:2280734 reverse
c1.0001.2210	0	2,820146377	0,01239	4,935141464	1,5E-05	4,572346153	2E-06	c1.0001.2210 oxidoreductase 2283156:2284307 reverse
c1.0001.2248	0	3,925297881	0	7,465269613	9,4E-19	7,103615189	4E-32	c1.0001.2248 hypothetical protein ABAZ39
c1.0001.2255	K02517	5,531313429	0,04704	10,63959726	0,00067	10,24313742	4E-18	c1.0001.2255 lauroyl acyltransferase 2342947:2343855 forward
c1.0001.2277	K17865	2,437359534	0,00026	4,535921474	1,1E-09	4,315005462	1E-13	c1.0001.2277 MaoC family dehydratase 2365493:2365942 forward
c1.0001.2303	K01485	5,405411225	0,01604	11,37106721	1,8E-05	10,449471	1E-05	c1.0001.2303 amidohydrolase 2390228:2391475 reverse
c1.0001.2420	0	7,917935543	0	15,54442366	3,5E-31	14,85354876	2E-37	c1.0001.2420 Hypothetical Protein 2508829:2510682 forward
c1.0001.2729	0	2,017913076	0,00189	3,611968026	5,6E-09	3,415576061	1E-08	c1.0001.2729 hypothetical protein ABAZ39
c1.0001.2756	0	19,23401866	0	35,16299832	1,1E-15	32,46089002	3E-43	c1.0001.2756 conserved protein of unknown function (Usg-like) 2841879:2842142 forward
c1.0001.2757	K01185	5,906110028	0,00059	11,24498245	5E-13	10,77712485	7E-26	c1.0001.2757 muraminidase 2842278:2842769 forward
c1.0001.2796	K07483	3,270053465	0,01463	6,210559406	0,00363	5,715836846	0,0025	c1.0001.2796 transposase 2888031:2888453 forward
c1.0001.2832	0	2,089489484	0,00022	3,942264716	8,6E-09	3,722692356	5E-12	c1.0001.2832 Hypothetical Protein 2918349:2919635 reverse
c1.0001.2835	0	#DIV/0!	0,00057	Inf	0,00857	39,20255483	0,0242	c1.0001.2835 Hypothetical Protein 2921488:2921580 reverse
c1.0001.2852	0	4,356990152	0,02042	7,492741048	0,00378	6,354173258	0,0017	c1.0001.2852 Hypothetical Protein 2940884:2941525 reverse
P1.0002.0053	K02014	5,794833053	0,00072	11,37493943	0,00026	10,63963712	5E-17	P1.0002.0053 TonB-dependent receptor 3068849:3071332 forward
P1.0002.0063	K02030	3,333580145	2,1E-06	6,497004574	7,1E-12	6,157127304	2E-17	P1.0002.0063 amino acid ABC transporter substrate-binding protein 3079684:3080475 forward
P1.0002.0064	K02029	4,117378411	0,00548	7,169464668	0,00031	6,384930309	5E-05	P1.0002.0064 polar amino acid ABC transporter permease 3080555:3081226 forward
P1.0002.0065	K02029	8,668524014	3,5E-05	18,01479484	2E-07	16,11502462	7E-08	P1.0002.0065 amino acid ABC transporter permease 3081223:3081882 forward
P1.0002.0066	K02028	3,543386584	0,00976	7,251870583	4,9E-07	7,002298552	1E-07	P1.0002.0066 amino acid ABC transporter ATP-binding protein

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
								3081860:3082594 forward
P1.0002.0275	0	4,735081097	0,04483	9,164594627	0,00409	7,836301112	0,0006	P1.0002.0275 sorbitol dehydrogenase 3327631:3328635 reverse
P1.0002.0382	0	#DIV/0!	0,00455	Inf	0,0012	53,44824961	0,0041	P1.0002.0382 Hypothetical Protein 3421781:3421912 forward
P1.0002.0396	K21416	3,873938166	0,00898	7,190473872	0,00585	6,578173411	2E-09	P1.0002.0396 acetoin:2,6-dichlorophenolindophenol oxidoreductase subunit alpha 3433617:3434600 forward
P1.0002.0412	0	6,885275525	0,00269	13,15417736	0,00307	10,73359427	0,0006	P1.0002.0412 hypothetical protein 3448699:3449082 forward
P1.0002.0433	0	2,774265554	5,3E-15	5,250922254	2,8E-14	4,989339471	3E-21	P1.0002.0433 hypothetical protein 3467208:3467588 reverse
P1.0002.0488	0	4,130389756	0,00162	7,827920977	0,00364	7,056436929	0,0008	P1.0002.0488 transposase 3525514:3525768 reverse
P1.0002.0583	K16841	27,714979	1,6E-08	53,09088109	4,6E-12	42,66723751	3E-12	P1.0002.0583 Asp/Glu/hydantoin racemase 3620205:3620951 reverse
P1.0002.0584	K02053	26,49784072	0,01076	49,77167942	1,2E-05	41,39141241	2E-10	P1.0002.0584 ABC transporter permease 3620969:3621808 reverse
P1.0002.0585	K02054	16,16397975	5,7E-06	29,08695426	3E-23	26,94436217	4E-35	P1.0002.0585 ABC transporter permease 3621812:3622735 reverse
P1.0002.0586	K02055	44,38656052	0	75,56340385	4,6E-51	67,65029837	4E-60	P1.0002.0586 ABC transporter substrate-binding protein 3622907:3624223 reverse
P1.0002.0587	K02052	6,578281233	0,00258	11,00986113	1,1E-05	9,65666336	3E-06	P1.0002.0587 spermidine/putrescine ABC transporter ATPase 3624289:3625308 reverse
P1.0002.0631	K01704	2,940048371	0,00439	5,583936938	0,00011	5,129148247	2E-05	P1.0002.0631 isopropylmalate isomerase 3668462:3669058 forward
P1.0002.0681	0	2,84157465	0,04842	5,316354914	0,00091	4,802236176	0,0003	P1.0002.0681 membrane protein 3717473:3718390 reverse
P1.0002.0720	K11717	2,937569955	2,6E-11	5,422630745	1,3E-13	5,116327661	1E-17	P1.0002.0720 cysteine desulfurase 3758592:3759854 reverse
P1.0002.0723	K09014	2,175885967	1,4E-07	4,086939399	8,4E-11	3,887342306	7E-17	P1.0002.0723 cysteine desulfarase 3762012:3763499 reverse
P1.0002.0724	0	3,589208505	2,2E-10	6,731116224	2,9E-10	6,329837357	6E-14	P1.0002.0724 Rrf2 family transcriptional regulator 3763548:3764030 reverse
P1.0002.0803	K07712	4,965298655	1,2E-09	9,309956325	1,4E-23	8,855594509	2E-32	P1.0002.0803 ntrC 3853490:3854935 reverse
P1.0002.0804	K07708	3,791420851	2,5E-06	7,025863842	2,3E-18	6,677411243	1E-23	P1.0002.0804 TRB
P1.0002.0805	K05540	2,289717944	0,03137	4,384328164	8,3E-11	4,204196254	9E-13	P1.0002.0805 US
P1.0002.0927	0	7,810973031	0,01612	14,31677796	0,01132	12,30029471	1E-06	P1.0002.0927 methyltransferase type 12 3984244:3985011 forward
P1.0002.0936	0	2,588692508	0,03307	4,593755234	4,7E-06	4,348211608	5E-08	P1.0002.0936 hypothetical protein 3991726:3992337 reverse
P1.0002.0937	0	4,180240455	3,6E-08	7,450101856	1,9E-10	6,917793695	2E-13	P1.0002.0937 hypothetical protein ABAZ39
P1.0002.0976	K06978	4,277594624	0,00062	7,825907407	2,1E-09	7,24451465	2E-12	P1.0002.0976 X-Pro dipeptidyl-peptidase 4037037:4038710 forward

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
P1.0002.0989	0	3,199915996	0,00015	5,674174617	9,2E-07	5,290723582	3E-08	P1.0002.0989 Hypothetical Protein 4049805:4050095 forward
P1.0002.1014	K01524	2,265672178	0	4,125043618	1,2E-10	3,916787263	6E-14	P1.0002.1014 exopolyphosphatase 4079474:4080502 reverse
P1.0002.1132	K04761	4,103332594	0,01752	7,61178505	9,6E-06	7,279717223	9E-19	P1.0002.1132 LysR family transcriptional regulator 4206348:4207304 forward
P1.0002.1150	K01486	2,119014149	0,02226	4,012402117	8,3E-11	3,827706892	2E-15	P1.0002.1150 adenine deaminase 4222104:4223813 reverse
P1.0002.1151	0	3,23882035	0	5,960266072	8,1E-16	5,636808528	5E-21	P1.0002.1151 exported protein of unknown function 4223848:4224177 reverse
P1.0002.1464	0	3,086896486	0,00545	5,33947151	0,00049	4,764252291	0,0003	P1.0002.1464 hypothetical protein GMDG
P1.0002.1541	K00266	5,756552347	5,3E-05	10,97580799	2,7E-22	10,47883085	4E-36	P1.0002.1541 LTD
P1.0002.1542	K00265	8,9837641	0	16,78994557	3,9E-16	15,96967403	3E-60	P1.0002.1542 Chain F, The 9.5 A Resolution Structure Of Glutamate Synthase From Cryo-Electron Microscopy And Its Oligomerization Behavior In Solution: Functional Implications 4641605:4645996 forward
P1.0002.1603	K00362	20,86481429	0,00017	32,81912563	5,8E-14	26,88824339	2E-16	P1.0002.1603 nitrite reductase 4701518:4703983 reverse
P1.0002.1604	K15576	15,04017778	0,00066	28,39634544	6,3E-11	26,59160578	2E-23	P1.0002.1604 nitrate ABC transporter substrate-binding protein 4704629:4706029 forward
P1.0002.1625	0	4,034945754	2,2E-13	7,428297841	1,1E-09	6,910761782	9E-13	P1.0002.1625 hypothetical protein 4730637:4730939 forward
P2.0003.0022	0	2,111072438	0,04567	3,837890972	8,7E-07	3,650602777	3E-08	P2.0003.0022 arabinose ABC transporter permease 4792049:4793356 reverse
P2.0003.0048	K03088	4,784930879	0	9,211262786	2,3E-19	8,747349462	2E-25	P2.0003.0048 RNA polymerase sigma factor 4810799:4811389 reverse
P2.0003.0063	K06016	4,173702765	0,00031	7,517596587	3,1E-08	6,84300251	1E-09	P2.0003.0063 allantoate amidohydrolase 4825623:4826852 reverse
P2.0003.0065	K02032	4,436250418	1,9E-06	8,553495741	1,1E-20	8,186683643	4E-25	P2.0003.0065 ABC transporter ATP-binding protein 4827841:4829505 reverse
P2.0003.0066	K02034	7,023978631	7E-05	13,55742118	1,1E-19	12,99889597	4E-32	P2.0003.0066 ABC transporter permease 4829505:4830377 reverse
P2.0003.0067	K02033	7,452305748	0	13,89300455	6,1E-29	13,17395865	5E-48	P2.0003.0067 ABC transporter permease 4830374:4831354 reverse
P2.0003.0068	K02035	5,188817309	2E-15	9,938243981	2,9E-18	9,465881142	2E-39	P2.0003.0068 peptide ABC transporter substrate-binding protein 4831487:4833097 reverse
P2.0003.0069	K01438	2,837959172	7,6E-05	5,258092664	8E-14	5,000953121	1E-17	P2.0003.0069 acetylornithine deacetylase 4833139:4834413 reverse
P2.0003.0092	0	5,29623871	0,00017	10,88989348	6,6E-09	10,30638171	4E-10	P2.0003.0092 putative two component transcriptional regulator

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
								4861388:4862161 reverse
P2.0003.0101	0	65,60259681	0	103,2846577	4,6E-17	76,36901615	2E-19	P2.0003.0101 methyltransferase, UbiE/COQ5 family 4872864:4873478 forward
P2.0003.0102	0	100,1843879	4,4E-16	174,0649324	1,5E-67	156,2512314	9E-83	P2.0003.0102 hypothetical protein; putative signal peptide 4873520:4874485 reverse
P2.0003.0103	0	251,6254103	1,8E-12	478,1694722	5E-59	387,8392709	8E-59	P2.0003.0103 hypothetical protein ABAZ39
P2.0003.0108	K02053	7,902687625	0,04227	13,26668034	0,00153	12,11985591	6E-09	P2.0003.0108 ABC transporter permease 4881268:4882083 reverse
P2.0003.0109	K02054	7,159374068	4,6E-07	14,27373069	7,2E-12	13,46243975	4E-15	P2.0003.0109 ABC transporter permease 4882076:4882966 reverse
P2.0003.0110	K02055	66,09104717	0	119,5256512	3,4E-28	98,66331378	1E-31	P2.0003.0110 ABC transporter substrate-binding protein 4883074:4884120 reverse
P2.0003.0111	K02052	10,49312965	2,5E-05	18,86334134	4,6E-10	16,67106315	4E-12	P2.0003.0111 spermidine/putrescine ABC transporter ATPase 4884167:4885330 reverse
P2.0003.0116	0	8,921354659	7E-06	15,56184306	7E-10	13,73775863	5E-12	P2.0003.0116 ABC transporter substrate-binding protein 4888149:4889162 forward
P2.0003.0117	K02050	22,28068367	6,4E-10	44,55794343	1,2E-14	39,42498084	1E-15	P2.0003.0117 ABC transporter permease 4889319:4890125 forward
P2.0003.0118	K02049	47,36140987	4,2E-05	73,34913232	3,8E-09	48,62355603	3E-08	P2.0003.0118 ABC transporter ATP-binding protein 4890129:4890911 forward
P2.0003.0119	K12960	#DIV/0!	0,0287	Inf	0,00045	61,35175086	0,0017	P2.0003.0119 hydrolase 4890908:4892239 forward
P2.0003.0120	K01799	#DIV/0!	0,00139	Inf	6,6E-05	75,79436851	0,0009	P2.0003.0120 Asp/Glu/hydantoin racemase 4892257:4892988 forward
P2.0003.0136	0	3,656489813	0,01069	6,84489563	0,00018	6,276741615	5E-05	P2.0003.0136 riboflavin biosynthesis protein RibD 4929097:4929885 forward
P2.0003.0158	K01992	#DIV/0!	0,01908	Inf	0,01715	47,37193095	0,0104	P2.0003.0158 multidrug ABC transporter permease 4950071:4950886 forward
P2.0003.0164	0	6,830183201	0,03675	10,52531676	0,01492	8,34129997	0,0222	P2.0003.0164 aspartate/glutamate/uridylate kinase 4954269:4954910 reverse
P2.0003.0266	K06200	6,841911858	0	12,9159598	7,3E-06	12,28995117	5E-53	P2.0003.0266 carbon starvation protein CstA 5075972:5077783 forward
P2.0003.0390	K02035	3,867165257	0,00025	7,521047653	1E-10	7,192880013	2E-13	P2.0003.0390 putative peptide transporter subunit: periplasmic-binding component of ABC superfamily 5226994:5228496 forward
P2.0003.0393	K02031	3,684083673	0,03729	7,215355748	4E-06	6,968409024	9E-07	P2.0003.0393 peptide ABC transporter ATP-binding protein

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
								5230405:5232528 forward
P2.0003.0438	K17323	7,399150622	0,01812	13,93083799	0,00135	11,77780999	0,0005	P2.0003.0438 sugar ABC transporter permease 5278691:5279488 reverse
P2.0003.0441	K17324	6,117376706	0,01534	9,60946593	4,8E-05	8,126650106	4E-05	P2.0003.0441 ABC transporter ATP-binding protein 5281493:5282596 reverse
P2.0003.0442	K00111	7,855449518	0,00147	13,93424516	5,3E-08	12,22434598	3E-09	P2.0003.0442 glycerol-3-phosphate dehydrogenase 5282677:5284218 reverse
P2.0003.0469	0	2,009468593	0,00037	3,604409188	1,2E-08	3,416305486	5E-09	P2.0003.0469 hypothetical protein ABAZ39
P2.0003.0470	0	3,707753889	0,00016	6,914013741	3,3E-18	6,587614301	1E-22	P2.0003.0470 hypothetical protein 5325726:5326022 reverse
P2.0003.0579	K06950	2,313730027	0,00198	4,490203658	3,5E-12	4,277641328	3E-15	P2.0003.0579 hydrolase 5449129:5449812 forward
P2.0003.0659	K02030	4,806049777	1,3E-12	8,919890857	1E-14	8,405108748	6E-21	P2.0003.0659 ABC transporter permease 5533586:5534383 reverse
P2.0003.0670	K07814	5,574008888	0	10,65693071	7,7E-19	10,13589522	9E-42	P2.0003.0670 divalent ion tolerance protein CutA 5548333:5549400 reverse
P2.0003.0672	0	2,016695029	0,00659	3,611675627	1,8E-08	3,42210233	6E-09	P2.0003.0672 exported protein of unknown function 5550184:5550864 forward
P3.0004.0030	0	2,009806341	0,00141	3,603432316	2,9E-09	3,4156531	2E-09	P3.0004.0030 hypothetical protein ABAZ39
P3.0004.0061	K10210	4,539910831	0,01669	8,048848971	1,1E-05	7,291871403	2E-06	P3.0004.0061 phytoene desaturase 5644645:5646186 reverse
P3.0004.0119	0	2,712532918	0,01546	4,9725546	0,00724	4,494343265	5E-05	P3.0004.0119 Hypothetical Protein 5707916:5708182 reverse
P3.0004.0124	0	3,813763878	0,02722	6,44638119	0,00942	5,486019621	0,0028	P3.0004.0124 hypothetical protein 5710683:5711120 forward
P3.0004.0152	0	2,655619214	0,00038	4,843950904	3,5E-13	4,60622496	7E-17	P3.0004.0152 Hypothetical Protein 5743925:5744713 reverse
P3.0004.0156	0	7,492680335	0,02073	13,28435851	0,00024	11,52229302	0,0002	P3.0004.0156 membrane protein 5746939:5747982 forward
P3.0004.0162	0	2,108064681	0,01832	3,770844259	0,00671	3,447360682	0,0029	P3.0004.0162 hypothetical protein 5751740:5751964 reverse
P3.0004.0179	K05800	14,50852096	6,1E-05	28,41310967	2,1E-05	26,18002136	5E-23	P3.0004.0179 ArsR family transcriptional regulator 5771479:5771958 reverse
P3.0004.0181	K01738	11,45035752	0,00137	22,1675275	0,00042	18,99488663	3E-08	P3.0004.0181 cysteine synthase 5772540:5773691 forward
P3.0004.0182	0	13,48795033	0,0013	24,00025846	5,1E-08	20,52031651	1E-08	P3.0004.0182 hypothetical protein 5773767:5774264 forward
P3.0004.0219	K01465	26,83648096	0,00054	58,97100554	1,9E-08	44,87033004	4E-08	P3.0004.0219 dihydroorotase 5811811:5813082 reverse
P3.0004.0220	0	5,934988632	0,0017	10,77465548	2,2E-05	9,347299437	5E-06	P3.0004.0220 membrane protein 5813776:5814612 forward
P3.0004.0222	K01999	33,73113308	0	60,88477757	4,8E-28	53,32489541	3E-33	P3.0004.0222 ABC transporter substrate-binding protein 5816254:5817408 forward
P3.0004.0223	K01997	13,05995985	0,00077	25,46303739	8,2E-10	23,49697045	5E-13	P3.0004.0223 ABC transporter permease 5817640:5818512 forward
P3.0004.0224	K01998	7,548937102	0,00062	14,86309816	3,1E-07	13,5154673	1E-08	P3.0004.0224 ABC transporter permease 5818509:5819546

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
								forward
P3.0004.0225	K01995	16,87449356	0,00125	36,8872134	1,3E-05	28,51975639	1E-05	P3.0004.0225 ABC transporter 5819543:5820361 forward
P3.0004.0226	K01996	7,383486133	0,00015	13,61681073	2,3E-06	12,00185176	1E-07	P3.0004.0226 amino acid ABC transporter ATPase 5820348:5821052 forward
P3.0004.0227	K03520	4,750126088	0,00659	9,056188332	2,4E-07	8,677952172	5E-11	P3.0004.0227 carbon monoxide dehydrogenase 5821074:5823443 forward
P3.0004.0294	K03797	2,014483054	0,01062	3,783557359	6,7E-07	3,590912477	6E-10	P3.0004.0294 peptidase S41 5883299:5884516 reverse
P3.0004.0301	0	3,71454452	1,3E-08	7,171932705	1,2E-17	6,836842926	1E-27	P3.0004.0301 diguanylate cyclase 5891219:5893603 forward
P3.0004.0335	0	2,905601215	0,00112	5,396166676	3,4E-15	5,141641755	8E-20	P3.0004.0335 Hypothetical Protein 5929771:5930478 forward
P3.0004.0336	0	5,361823869	7,5E-10	9,988989753	3,9E-13	9,405846817	4E-17	P3.0004.0336 Hypothetical Protein 5930493:5930762 forward
P3.0004.0337	0	2,829709834	6,8E-05	5,460642247	1E-12	5,200452806	1E-16	P3.0004.0337 Hypothetical Protein 5930793:5932055 forward
P3.0004.0338	0	3,204464711	1,2E-06	6,158282007	8,3E-15	5,880013677	5E-19	P3.0004.0338 Hypothetical Protein 5932241:5933518 forward
P3.0004.0379	0	2,191257165	1,7E-05	4,11310816	3,6E-10	3,898455455	7E-14	P3.0004.0379 hypothetical protein ABAZ39
P3.0004.0411	0	4,385243649	2,2E-16	7,684210173	2,6E-18	7,279003505	5E-15	P3.0004.0411 Hypothetical Protein 6014203:6014274 reverse
P3.0004.0416	0	2,371872773	0,00454	4,340702399	4,9E-08	4,121369946	2E-10	P3.0004.0416 hypothetical protein ABAZ39
P3.0004.0423	K03458	3,00216349	0,00364	5,688243825	1,7E-08	5,45143403	4E-10	P3.0004.0423 purine permease 6026596:6027921 reverse
P3.0004.0430	K02055	2,212785728	0,03492	4,234544732	3,5E-10	4,060297928	3E-12	P3.0004.0430 ABC transporter substrate-binding protein 6034164:6035285 reverse
P3.0004.0472	K02035	5,537401122	0,00016	10,5247041	1E-17	10,05354631	1E-32	P3.0004.0472 putative oligopeptide ABC transporter (substrate-binding protein) 6083567:6085219 forward
P3.0004.0473	K02033	12,59050553	0,00014	24,3736014	7,8E-16	23,36728166	1E-36	P3.0004.0473 ABC transporter permease 6085371:6086348 forward
P3.0004.0474	K02034	4,228810447	0,0128	7,825795874	0,00016	7,191018909	8E-05	P3.0004.0474 peptide ABC transporter permease 6086345:6087247 forward
P3.0004.0475	K02031	4,480310813	0,0014	8,37150013	1,3E-06	7,820257208	8E-08	P3.0004.0475 methionine ABC transporter ATP-binding protein 6087261:6088250 forward
P3.0004.0522	K10547	3,656489813	0,03562	6,84489563	0,00018	6,276733579	5E-05	P3.0004.0522 ABC transporter permease 6136149:6137360 forward
P4.0005.0126	0	#DIV/0!	7E-08	Inf	6,7E-05	77,00204	0,0006	P4.0005.0126 Hypothetical Protein 6348018:6348215 forward
P5.0006.0092	0	4,206554894	6,2E-10	7,672045519	1E-08	7,115462986	3E-17	P5.0006.0092 chromosome partitioning protein ParB 6544577:6545446 forward
P6.0001.0015	0	3,895924656	0,01305	7,147750104	0,00032	6,394706498	5E-05	P6.0001.0015 hypothetical protein APCd
P6.0001.0017	0	5,686429379	0,01357	9,940693507	0,00092	8,406190692	0,0002	P6.0001.0017 hypothetical protein APCd
P6.0001.0020	0	10,02033853	0,00153	15,62393265	0,03137	10,44331732	0,02	P6.0001.0020 hypothetical protein APCd

ID	KO	Fold Change CLC	P-value CLC	Fold_Change DESeq	P-value DESeq	Fold_Change EdgeR	P-value EdgeR	Function
P6.0001.0022	0	5,951343849	2,2E-16	11,41048678	1,7E-10	10,54358241	4E-13	P6.0001.0022 hypothetical protein APCd
P6.0001.0023	0	2,782509608	0,00268	5,031606749	0,00153	4,670656008	0,0005	P6.0001.0023 hypothetical protein APCd
P6.0001.0026	0	6,05836567	7,5E-08	10,87817727	1,1E-05	9,66972181	1E-06	P6.0001.0026 hypothetical protein APCd
P6.0006.0004	K19591	2,270957092	0,0033	4,183381277	0,00015	3,847195726	2E-05	P6.0006.0004 MerR family transcriptional regulator 140421:140852 reverse
P6.0006.0013	0	2,059887971	0,0001	3,896469118	1,1E-06	3,686739366	2E-09	P6.0006.0013 short-chain dehydrogenase 148234:148398 forward
P6.0007.0012	0	6,885275525	0,00048	13,15417736	0,00307	10,73351801	0,0006	P6.0007.0012 Hypothetical Protein 169889:170146 forward
P6.0014.0003	0	2,542877052	0,01645	4,545433973	7,7E-05	4,173265848	9E-06	P6.0014.0003 transcriptional regulator 242769:243674 forward
P6.0014.0004	K00101	2,868649369	0,02139	5,462966832	2,5E-05	5,155157196	1E-06	P6.0014.0004 lactate dehydrogenase 243787:244935 reverse
P6.0016.0007	0	8,596436865	0,00259	17,40245853	9,3E-06	15,04332323	7E-06	P6.0016.0007 alpha/beta hydrolase 254653:255645 forward
P6.0019.0006	0	3,08259264	0,00175	5,887318773	0,00076	5,384410553	6E-06	P6.0019.0006 hypothetical protein 318275:318511 reverse
P6.0021.0011	0	6,711398629	0,0422	12,68000569	0,00304	10,65821694	0,0007	P6.0021.0011 ABC transporter permease 330790:331725 forward
P6.0023.0016	0	4,63014454	0,03125	7,188309545	0,0019	6,208584161	0,0035	P6.0023.0016 Hypothetical Protein 360807:360869 reverse
P6.0030.0008	0	4,599165487	0,00036	8,613659034	1,1E-06	7,889214827	9E-11	P6.0030.0008 hypothetical protein 415910:417328 reverse
P6.0041.0007	K18954	7,060250418	0,01002	11,03961188	0,01511	8,406005076	0,0178	P6.0041.0007 AraC family transcriptional regulator 495893:496222 reverse
P6.0046.0005	K02050	8,086902615	0,02935	15,18167029	0,0007	12,87884032	0,0003	P6.0046.0005 nitrate ABC transporter permease 528163:528900 forward
P6.0051.0007	K07675	2,714303458	0,01334	5,258994047	6,8E-05	4,886228389	5E-07	P6.0051.0007 Hypothetical Protein 551243:552406 reverse
P6.0056.0001	K00799	8,086902615	0,0293	15,18167029	0,0007	12,87880695	0,0003	P6.0056.0001 Hypothetical Protein 571251:571808 reverse
P6.0083.0001	0	2,223673625	0,02263	4,105132545	3E-10	3,912169017	2E-15	P6.0083.0001 Hypothetical Protein 631943:633355 forward
P6.0442.0001	0	2,088238413	0,00402	3,915082793	0,00186	3,713455463	3E-12	P6.0442.0001 hypothetical protein 718154:718375 reverse
P6.0486.0001	0	3,138661968	0,01805	6,305613692	0,00294	5,989585137	1E-12	P6.0486.0001 Hypothetical Protein 723935:724447 forward
P6.2445.0001	0	7,290317635	0,00545	11,553907	0,01528	8,507667677	0,0224	P6.2445.0001 Hypothetical Protein 950167:950367
P6.2523.0001	0	2,406144588	0,0444	4,417399066	0,00456	3,964037669	0,0014	P6.2523.0001 Hypothetical Protein 959202:959690 reverse
P6.3080.0001	0	#DIV/0!	1,5E-05	Inf	0,00321	46,5066641	0,0108	P6.3080.0001 Hypothetical Protein 1023327:1023503 forward

Anexo 03: Comparação dos genes diferencialmete expressos da *Azospirillum brasilense* FP2 com os genes diferencialmente expressos de *Azoarcus sp*. BH72, *Azotobacter vinelandii e Paenibacillus sp.* WLY78 em condições de Fixação Biológica de Nitrogênio.

Nome	foldchange *Azospirillum brasilense* FP2	foldchange *Azoarcus sp.* BH72	foldchange *Azotobacter vinelandii*	foldchange *Paenibacillus sp.* WLY78	Gene_product
nifD	587,90	22	53,1	307,29	Nitrogenase molybdenum-iron protein alpha chains, NifD
nifK	446,70	-	55,1	215,54	Nitrogenase molybdenum-iron protein beta chains, NifK
nifS	377,40	18	8	-	Nitrogen fixation cysteine desulfurase, NifS
nifX	369,47	4,9	2	213,81	Nitrogenase molybdenum-cofactor biosynthesis protein, NifX
nifN	339,30	21	3,8	158,94	Nitrogenase molybdenum-cofactor biosynthesis protein, NifN
nifU	323,52	-	13,1	-	Nitrogen fixation Fe-S cluster scaffold protein, NifU
nifH	288,65	-	143,8	446,18	Nitrogenase iron protein, NifH
fixB	237,97	-	40,5	-	Electron transfer flavoprotein, alpha subunit, FixB
nifW	225,90	7	4,9	-	Nitrogen fixation protein, NifW
fixA	201,28	-	8,5	-	Electron transfer flavoprotein beta-subunit, FixA
fixC	166,49	-	3,5	-	Electron-transferring-flavoprotein dehydrogenase, FixC
nifB	133,20	10	18	1039,39	Nitrogenase cofactor biosynthesis protein, NifB
nifE	112,80	16	6,3	228,02	Nitrogenase molybdenum-cofactor biosynthesis protein, NifE
amtB	108,80	-2,2	-	200,47	Ammonia permease
nifP	108,80	3,1	-	-	serine acetyltransferase
nifV	62,09	3	6,7	372,62	Nitrogen fixation homocitrate synthase, NifV
fixX	51,77	-	16,2	-	FixX ferredoxin-like protein
fdxB	50,87	8,7	-	-	ferredoxin III (nif specific)
nifZ	37,64	14	4,9	-	Nitrogen fixation protein, NifZ
nirB	33,65	-2,1	-	3,02	Nitric oxidereductaseflavoprotein
glnP	18,22	2	-	-	glutamine ABC transporter; permease component
gltB	16,96	-	-	5,97	Glutamate synthase domain 2
fdxN	16,96	37	-	-	ferredoxin

Nome	foldchange *Azospirillum brasilense* FP2	foldchange *Azoarcus sp.* BH72	foldchange *Azotobacter vinelandii*	foldchange *Paenibacillus sp.* WLY78	Gene_product
nifQ	16,58	11	12	-	Nitrogen fixation cofactor assembly protein, NifQ
gltD	11,05	-	-	14,72	NADPH-dependent glutamate synthase beta chain
glnH	6,58	2,9	-	-	glutamine ABC transporter; substrate-binding component
nifA	5,49	1,8	0,9	-	Nif-specific ?54-dependent transcriptional activator protein, NifA
sufS	5,48	-	-	7,47	Cysteine desulfurase
copA	3,19	-2,5	-	-	copper-transporting P-type ATPase
rpoH	2,99	-3	-	-	RNA polymerase sigma factor (sigma32)
fabI	2,90	-	-	1,73	Enoyl-[acyl-carrier-protein] reductase (NADH)
atpH	-2,04	-	1,4	10,11	F0F1-type ATP synthase, delta subunit
rplL	-2,16	2,1	-	-	50S ribosomal subunit protein L7/L12
rpoC	-2,23	-3,3	-	-	DNA-directed RNA polymerase subunit beta
sdhD	-2,48	-1,8	-	-	succinate dehydrogenase, hydrophobic membrane anchor subunit
lepA	-2,61	-2	-	-	GTP-binding elongation factor
prsA	-2,68	-1,8	-	-	ribose-phosphate pyrophosphokinase (RPPK)
trpE	-2,71	-3,4	-	-	anthranilate synthase
pdhA	-2,79	-1,8	-	-	pyruvate dehydrogenase E1 component, alpha subunit
engA	-2,84	-1,9	-	-	ribosome-associated GTPase
aroF	-2,87	1,8	-	-	3-deoxy-7-phosphoheptulonate synthase
petB	-3,14	-	0,8	-	Cytochrome b/b6
nuoA	-3,15	-	1,8	-	NADH-ubiquinone oxidoreductase, chain alpha
lpdA	-3,23	-1,8	-	-	dihydrolipoyl dehydrogenase
atpG	-3,34	-	1,1	2,26	F0F1-type ATP synthase, gamma subunit
dapA	-3,76	1,9	-	-	dihydrodipicolinate synthase
plsX	-3,88	-2,6	-	-	fatty acid/phospholipid synthesis protein
fabH	-3,89	-1,8	-	-	3-oxoacyl-[acyl-carrier-protein] synthase III
smpB	-3,98	2,1	-	-	trans-translation protein (ssrA RNA-binding protein)

Nome	foldchange *Azospirillum brasilense* FP2	foldchange *Azoarcus sp.* BH72	foldchange *Azotobacter vinelandii*	foldchange *Paenibacillus sp.* WLY78	Gene_product
hypB	-4,03	-	3,2	-	Hydrogenase nickel incorporation protein, HypB
rplO	-4,12	-4,2	-	-	50S ribosomal subunit protein L15
rplU	-4,14	1,8	-	-	50S ribosomal subunit protein L21
atpD	-4,37	3,2	1,6	0,96	F0F1-type ATP synthase, beta subunit
atpI	-4,39	-	5,4	20,76	F0F2-type ATP synthase, subunit I
nuoF	-4,41	-	1,4	-	NADH-quinone oxidoreductase, F subunit
petA	-4,42	-	1,1	-	Ubiquinol-cytochrome c reductase, iron-sulfur subunit
purB	-4,88	-1,9	-	-	Adenylosuccinate lyase
nuoI	-4,98	-	1,8	-	NADH-quinone oxidoreductase, chain I
hypC	-5,21	-	1,6	-	Hydrogenase assembly chaperone, HypC/HupF
petC	-5,34	-	1,9	-	Ubiquinol--cytochrome c reductase, cytochrome c1
nuoL	-5,45	-	1,4	-	NADH-plastoquinone oxidoreductase, chain L
tatB	-5,45	-1,8	-	-	TatB protein translocation system subunit
rpsG	-5,53	-2,3	-	-	30S ribosomal subunit protein S7
rpsQ	-5,69	-2,9	-	-	30S ribosomal subunit protein S17
nuoG	-5,88	-	0,8	-	NADH-quinone oxidoreductase, chain G
atpC	-5,92	-3	1	1,28	F0F1-type ATP synthase, epsilon subunit
rpsT	-6,52	2,3	-	-	30S ribosomal protein S20
rplE	-6,60	-3,1	-	-	50S ribosomal subunit protein L5
atpF	-6,63	-	1,6	11,85	F0F1-type ATP synthase, subunit B
rplX	-6,72	-4,3	-	-	50S ribosomal subunit protein L24
atpB	-6,81	-7,1	1,4	4,77	F0F1-type ATP synthase, subunit A
rpsM	-6,81	-5,1	-	-	30S ribosomal subunit protein S13
hemH	-6,96	-	-	3,38	Protohemeferro-lyase (ferrochelatase)
nuoM	-7,01	-	1,1	-	Proton-translocating NADH-quinone oxidoreductase, chain M
tufB	-7,31	-8,9	-	-	elongation factor Tu (EF-Tu)

Nome	foldchange *Azospirillum brasilense* FP2	foldchange *Azoarcus sp.* BH72	foldchange *Azotobacter vinelandii*	foldchange *Paenibacillus sp.* WLY78	Gene_product
pdhB	-7,36	-1,8	-	-	pyruvate dehydrogenase E1 component, beta subunit
cysC	-8,08	-	-	29,07	Adenylylsulfate kinase
rpsR	-8,81	-2,8	-	-	30S ribosomal subunit protein S18
nuoJ	-9,31	-	1,4	-	NADH-ubiquinone/plastoquinone oxidoreductase, chain J
atpE	-9,86	-	1,5	7,38	F0F1-type ATP synthase, subunit C
nuoK	-10,26	-	1,6	-	NADH-ubiquinone oxidoreductase, chain K
rplP	-11,92	2,4	-	-	50S ribosomal subunit protein L29
rplV	-14,12	-4,1	-	-	50S ribosomal subunit protein L22
cysH	-14,45	-	-	9,56	Phosphoadenosinephosphosulfate reductase
efp	-17,08	1,8	-	-	elongation factor P (EF-P)
rplI	-21,73	-2,7	-	-	50S ribosomal protein L9
rpmC	-22,81	6,1	-	-	50S ribosomal subunit protein L29
gdhA	-25,80	-	-	12,22	Glutamate dehydrogenase/leucine dehydroGenase
rplJ	-32,35	4,4	-	-	50S ribosomal subunit protein L10

Anexo 04: Comparação dos dados de RNA-Seq com dados de proteômica da *A. brasilense* FP2.

Tabela 01: genes diferencialmente expressos em *A. brasilense* FP2 que codificam para proteínas que também foram identificadas no proteoma em condição de Fixação Biológica de Nitrogênio (-N). Os valores de log2FoldChange são do transcriptoma analisados pelo *software* DESeq, em tons de verde os genes ativados e em tons de vermelho os genes reprimidos. As células em tons de cinza são spots que foram identificados no mínimo duas vezes com a mesma proteína, sendo selecionado a linha com maior valor de volume.

Nome	log2FoldChange DESeq	Protein_BLAST	Vol.
nifD	9,199	nitrogenase molybdenum-iron protein alpha chain, nifD	0,35
nifK	8,803	nitrogenase molybdenum-iron protein beta chain, nifK	0,35
nifU	8,3377	Fe-S cluster assembly protein	0,319
nifH	8,173	nitrogenase iron protein, nifH, dinitrogenase reductase	1,772
fixX	7,379	4Fe-4S ferredoxin	0,069
	7,197	conserved protein of unknown function	0,098
fixC	5,693	4Fe-4S ferredoxin	0,069
	5,597	short-chain amide or branched-chain amino acid ABC transporter, ATP-binding component	0,043
	3,435	conserved protein of unknown function, putative Thioredoxin domain	0,102
	2,096	putative spermidine/putrescine ABC transporter (Substrate binding protein)	0,194
	1,639	conserved protein of unknown function	0,147
sufC	1,513	ABC transporter involved in Fe-S cluster assembly SufC, ATP-binding protein	0,074
hspD1	-1,029	small heat shock protein, HSP20-like chaperone	0,039
rplL	-1,11	50S ribosomal subunit protein L7/L12	1,385
acnB	-1,215	aconitate hydratase 2 [EC:4.2.1.3]	0,496
purA	-1,379	Adenylosuccinate synthetase	0,127
accB	-1,458	acetyl-CoA carboxylase, biotin carboxyl carrier subunit	0,098
bcp ou glnE	-1,504	thiol peroxidase, thioredoxin-dependent	0,066

Nome	log2FoldChange DESeq	Protein_BLAST	Vol.
sdhB ou sdhA	-1,51	succinate dehydrogenase, Fe-S subunit	0,146
	-1,526	conserved protein of unknown function, exopolyphosphatase domain	0,102
	-2,017	putative peptidyl-prolyl isomerase (PpiC-like)	0,008
petA	-2,143	Ubiquinol-cytochrome C reductase, iron-sulfur subunit	0,175
pyrE2	-2,159	orotate phosphoribosyltransferase	0,087
pyrG	-2,252	CTP synthase	0,113
rplI ou	-2,532	50S ribosomal protein L9	0,349
rplE	-2,721	50S ribosomal subunit protein L5	0,174
rplS	-3,226	50S ribosomal protein L19	0,061
rpsC	-3,318	30S ribosomal subunit protein S3	0
efp	-4,094	elongation factor P (EF-P)	0,482
rplI ou	-4,441	50S ribosomal protein L9	0,349
hbdA	-4,51732	3-hydroxybutyryl-CoA dehydrogenase	0,065
rplL	-4,809	50S ribosomal subunit protein L7/L12	1,385

Tabela 02: genes diferencialmente expressos em *A. brasilense* FP2 que codificam para proteínas que também foram identificadas no proteoma em condição amoniotrófica (+N). Os valores de log2FoldChange são do transcriptoma analisados pelo *software* DESeq, em tons de verde os genes ativados e em tons de vermelho os genes reprimidos. As células em tons de cinza são *spots* que foram identificados no mínimo duas vezes com a mesma proteína, sendo selecionado a linha com maior valor de volume.

Nome	log2foldchange DESeq	Protein_BLAST	Vol.
lysS	2,3377	lysyl-tRNA synthetase	0,139
ilvE	1,719	branched-chain-amino-acid aminotransferase	0,289
pta2	1,716	phosphate acetyltransferase	0,035
tal	-1,056	transaldolase	0,108
fba	-1,261	fructose-bisphosphate aldolase	0,516
sucD	-1,418	succinyl-CoA synthetase, alpha subunit,NAD(P)-binding	0,125
dnaA	-1,435	chromosomal replication initiator protein dnaA	0,042
prpE	-1,443	propionyl-CoA synthetase	0,08
gatA	-1,518	glutamyl-tRNA amidotransferase, subunit A	0,075
lpdA	-1,693	dihydrolipoyl dehydrogenase	0,227
argG	-1,77	argininosuccinate synthase	0,415
etfA	-1,772	electron transfer flavoprotein alpha-subunit	0,559
aceB	-1,77504	malate synthase G (fragment)	0,249
ppdK	-1,776	pyruvate phosphate dikinase	0,025
leuA	-1,785	2-isopropylmalate synthase	0,404
IDH1, IDH2, icd	-1,807	isocitrate dehydrogenase	0,649
asd	-1,808	aspartate-semialdehyde dehydrogenase	0,359
dld, LDHD	-1,835	putative D-lactate dehydrogenase	
gltX2	-1,88163	glutamyl-tRNA synthetase	

Nome	log2foldchange DESeq	Protein_BLAST	Vol.
pdxY	-1,925	pyridoxal/pyridoxamine kinase	0,092
engD	-2,026	GTP-dependent nucleic acid-binding protein	
cysK2	-2,168	cysteine synthase	0,198
dapD	-2,17977	tetrahydrodipicolinate N-succinyltransferase	0,764
ilvH	-2,27003	acetolactate synthase, small subunit	
ilvC	-2,643	Ketol-acid reductoisomerase	0,218
cbbT	-2,8145	transketolase (TK)	0,258
ndk	-3,291	nucleoside diphosphate kinase	0,916
leuC	-3,403	3-isopropylmalate dehydratase, large subunit	0,346
fixJ1	-4,233	two component transcriptional regulatory protein, LuxR family	0,103
tyrC	-4,504	cyclohexadienyl dehydrogenase	
gdhA	-4,689	NADP-specific glutamate dehydrogenase	0,235

Tabela 03: genes diferencialmente expressos em *A. brasilense* FP2 que codificam para proteínas que também foram identificadas no proteoma em condição amoniotrófica (+N) e em condição de Fixação Biológica de Nitrogênio, simultaneamente. Os valores de log2FoldChange são do transcriptoma analisados pelo *software* DESeq, em tons de vermelho estão os genes reprimidos. As células em tons de cinza são *spots* que foram identificados no mínimo duas vezes com a mesma proteína, sendo selecionado a linha com maior valor de volume. O volume das duas condições estão mostrados (Vol. +N e –N).

Nome	log2foldchange DESeq	Vol. (-N)	Funcão	Vol. (+N)
leuD	-2,21579	**0,105**	Amino Acid Metabolism, Valine, leucine and isoleucine biosynthesis	**0,073**
mdh	-2,465	**1,064**	Citrate cycle (TCA cycle)	**1,536**
leuA	-2,673	**0**		**0,063**
nusG	-2,059	**0,029**	Genetic Information Processing, Transcription, Transcription machinery	**0,058**
tufB1	-2,87082	**2,377**	Genetic Information Processing, Translation, Translation factors	**1,663**
rplD ou rplC	-3,917	**0,056**	Genetic Information Processing, Translation, Ribosome	**0,035**
rplD	-4,91	**0,056**	Genetic Information Processing, Translation, Ribosome	**0,035**
adk ou secY	-1,847	**0,124**	Nucleotide Metabolism, Purine metabolism	**0,122**
	-2,197	**0,069**	Genetic Information Processing, Translation proteins	**0,079**
fusA1	-2,896	**1,526**	Translation, Translation factors	**0,334**
	-1,10467	**0,075**	Protein folding and associated processing	**0,173**
fabF2	-1,176	**0,131**	Lipid Metabolism, Fatty acid biosynthesis	**0,2**
hss	-1,153	**0,169**	etabolism, Biosynthesis of Other Secondary Metabolites	**0,197**
atpC	-2,564	**0,235**	Energy Metabolism, Oxidative phosphorylation	**0,152**
atpD	-2,12896	**1,449**	Oxidative phosphorylation	**1,804**
idiA	-1,614	**0,1**	Environmental Information Processing, Membrane Transport, Transporters	**0**
rpsA	-1,431	**1,099**	Genetic Information Processing, Translation, Ribosome	**0,684**
pntAA1	-1,5768	**0,031**	Metabolism of Cofactors and Vitamins, Nicotinate and nicotinamide metabolism	**0,211**

Nome	log2foldchange DESeq	Vol. (-N)	Funcão	Vol. (+N)
phbC	-1,691	**0,254**	Butanoate metabolism	**0,469**
metY1	-2,098	**0**	Cysteine and methionine metabolism	**0,649**
	-2,052	**0,124**	Unclassified, Genetic Information Processing, Protein folding and associated processing	**0,177**
	-1,482	**0,269**	Unclassified, Cellular Processes and Signaling, Pores ion channels	**0,24**
	-1,482	**0,135**	Cellular Processes and Signaling, Pores ion channels	**0,24**
rpoA	-3,699	**1,271**	Genetic Information Processing, Translation, Translation factors	**0,319**
	-1,23017	**0**	Unclassified, Metabolism, Lipid metabolism	**0,189**
accC	-2,34	**0,094**	Fatty acid biosynthesis	**0,116**
trxB	-1,54291	**0,188**	Nucleotide Metabolism, Pyrimidine metabolism	**0,135**
gatB	-2,028	**0,072**	Genetic Information Processing, Translation, Aminoacyl-tRNA biosynthesis	**0,223**
metQ1	1,764	**0,361**	D-Methionine transport system	**0,358**
tyrS	-2,01925	**0**	Phenylalanine, tyrosine and tryptophan biosynthesis	**0,39**
	-2,255	**0,035**	transporter activity	**0,07**
	-1,09982	**0,099**	unknown function	**0,157**
sodB	-1,5191	**0,895**	Cellular Processes, Transport and Catabolism, Peroxisome	**0,736**
	-1,593	**0,076**	unknown function	**0,083**
	-1,13867	**0,139**	Environmental Information Processing, Membrane Transport, Transporters- Peptides/nickel transport system	**0,067**
proX	-1,461	**0,141**	Environmental Information Processing, Membrane Transport, ABC transporters	**2,208**
livK	-1,585	**0,427**	Membrane Transport, ABC transporters	**0,744**
rplY	-5,242	**0,539**	Genetic Information Processing, Translation, Ribosome	**0,716**
rlmN	-1,244	**0,047**	Genetic Information Processing, Translation, Ribosome biogenesis	**0,058**

Anexo 05: Tabelas com a relação dos novos genes adicionados ao genoma de *H. seropedicae* SMR1.

A Tabela 01 lista os novos genes anotados pelo NCBI. A Tabela 02 lista os novos genes anotados pelo HGF e no final da Tabela 02 há os genes anotados pelo Glimmer.

Tabela 01

Locus_tag_NCBI	Posição
HSERO_RS00070	complement (22849..23139)
HSERO_RS00160	complement (43614..43865)
HSERO_RS00190	50370..51365
HSERO_RS00215	complement (57620..57916)
HSERO_RS00220	complement (58008..58412)
HSERO_RS00245	complement (63109..63564)
HSERO_RS00260	66166..66453
HSERO_RS00355	complement (85184..86503)
HSERO_RS00385	complement (91531..92040)
HSERO_RS00790	166281..166535
HSERO_RS24055	201139..201369
HSERO_RS00955	203037..203966
HSERO_RS01025	complement (219133..220320)
HSERO_RS01075	complement (228196..228453)
HSERO_RS01410	305960..306370
HSERO_RS02125	464682..464882
HSERO_RS02275	complement (499310..499546)
HSERO_RS02490	545545..545823
HSERO_RS02550	complement (560502..560711)

Tabela 02

Locus_tag_HGF	Localização
SC_1_N_1	complement (5853..6221)
SC_1_N_9	complement (160736..161176)
SC_1_N_12	195716..195868
SC_1_N_20	351015..351185
SC_1_N_21	complement (357055..357207)
SC_1_N_23	complement (417380..417541)
SC_1_N_47	complement (671238..671342)
SC_1_N_51	787285..787434
SC_1_N_54	854002..854118
SC_1_N_55	855552..855704
SC_1_N_61	909224..909343
SC_1_N_62	909586..909642
SC_1_N_65	942587..942799
SC_1_N_73	1012582..1013079
SC_1_N_76	1025501..1025863
SC_1_N_77	complement (1037143..1037310)
SC_1_N_82	complement (1181939..1182055)
SC_1_N_84	1208197..1208334
SC_1_N_85	1236502..1236651

Locus_tag_NCBI	Posição
HSERO_RS02610	complement (571438..571836)
HSERO_RS02735	complement (607221..607640)
HSERO_RS02770	complement (612416..612664)
HSERO_RS02775	612920..613150
HSERO_RS02780	complement (613202..613435)
HSERO_RS02785	613837..614028
HSERO_RS02790	614126..614314
HSERO_RS03225	complement (709408..710982)
HSERO_RS03465	complement (755475..755663)
HSERO_RS03805	840613..840828
HSERO_RS03810	840947..841798
HSERO_RS03875	855828..856037
HSERO_RS03880	856099..856281
HSERO_RS03890	858330..858926
HSERO_RS04290	complement (941596..942048)
HSERO_RS04390	complement (973884..974237)
HSERO_RS04395	complement (974326..974532)
HSERO_RS04755	1041873..1042067
HSERO_RS04920	complement (1082207..1082554)
HSERO_RS04950	1089009..1089950
HSERO_RS05385	1189154..1189387
HSERO_RS05495	1244166..1245071
HSERO_RS05870	1335353..1335541
HSERO_RS06035	1364112..1364450
HSERO_RS06105	complement (1375511..1376029)
HSERO_RS06155	complement (1384308..1384538)
HSERO_RS06205	1397334..1397636
HSERO_RS06420	1450118..1450393

Locus_tag_HGF	Localização
SC_1_N_86	complement (1236679..1236861)
SC_1_N_88	1250204..1250533
SC_1_N_90	complement (1283931..1284101)
SC_1_N_92	1302505..1302759
SC_1_N_94	complement (1326891..1327037)
SC_1_N_99	complement (1365361..1365519)
SC_1_N_100	complement (1372628..1372993)
SC_1_N_105	1405382..1405504
SC_1_N_115	1569238..1569393
SC_1_N_116	complement (1597026..1597196)
SC_1_N_118	complement (1655791..1656192)
SC_1_N_119	complement (1689121..1689582)
SC_1_N_125	complement (1788554..1788787)
SC_1_N_129	1820006..1820116
SC_1_N_131	1867320..1867484
SC_1_N_140	1940814..1940990
SC_1_N_142	complement (1955902..1956096)
SC_1_N_146	complement (2022905..2023102)
SC_1_N_147	2026128..2026268
SC_1_N_150	2060930..2061013
SC_1_N_152	2078188..2078355
SC_1_N_153	2087564..2087761
SC_1_N_155	complement (2140584..2140727)
SC_1_N_158	2171618..2171782
SC_1_N_162	2222391..2222660
SC_1_N_165	2260132..2260248
SC_1_N_174	2378921..2379070
SC_1_N_180	2471343..2471597

Locus_tag_NCBI	Posição
HSERO_RS06505	1478650..1479987
HSERO_RS06655	1513196..1513426
HSERO_RS06730	complement (1530110..1530379)
HSERO_RS06745	1532624..1532839
HSERO_RS06850	1559045..1559254
HSERO_RS07095	complement (1615570..1616733)
HSERO_RS07270	1664194..1664373
HSERO_RS07790	complement (1784006..1784350)
HSERO_RS07795	1784349..1784774
HSERO_RS07830	complement (1791024..1791242)
HSERO_RS07890	1803590..1803811
HSERO_RS07940	1814615..1814986
HSERO_RS08250	complement (1880193..1880375)
HSERO_RS08260	1881824..1882096
HSERO_RS08390	1911677..1911961
HSERO_RS08485	1931540..1931785
HSERO_RS08600	complement (1956106..1956306)
HSERO_RS08700	complement (1973480..1973848)
HSERO_RS08810	1997434..1997772
HSERO_RS08975	2028471..2028923
HSERO_RS09115	2070736..2071221
HSERO_RS09120	2071223..2071504
HSERO_RS09375	2128754..2129104
HSERO_RS09490	complement (2158945..2159277)
HSERO_RS09805	2222667..2223035
HSERO_RS09835	2228170..2228391
HSERO_RS09875	complement (2233601..2233795)
HSERO_RS10035	complement (2271003..2271992)

Locus_tag_HGF	Localização
SC_1_N_196	2687406..2687570
SC_1_N_197	2693082..2693180
SC_1_N_198	2693184..2693384
SC_1_N_200	2695337..2695501
SC_1_N_201	2697585..2697695
SC_1_N_202	2697703..2697804
SC_1_N_206	2702031..2702192
SC_1_N_207	2702192..2702329
SC_1_N_208	complement (2703859..2704053)
SC_1_N_209	2710255..2710365
SC_1_N_216	complement (2804402..2804578)
SC_1_N_217	2811179..2811346
SC_1_N_222	2917238..2917393
SC_1_N_225	complement (2948506..2948988)
SC_1_N_226	complement (2949119..2949556)
SC_1_N_227	complement (2958269..2958343)
SC_1_N_258	complement (3308834..3308989)
SC_1_N_259	3326261..3326929
SC_1_N_267	complement (3426439..3426564)
SC_1_N_269	3515941..3516111
SC_1_N_270	complement (3528069..3528296)
SC_1_N_272	complement (3569487..3569720)
SC_1_N_278	complement (3626177..3626311)
SC_1_N_280	3630573..3630674
SC_1_N_286	3688561..3688641
SC_1_N_288	complement (3707821..3707964)
SC_1_N_289	complement (3708006..3708434)
SC_1_N_302	3912422..3913180

Locus_tag_NCBI	Posição
HSERO_RS10060	2276691..2276930
HSERO_RS10065	2277145..2277924
HSERO_RS10070	2278018..2278419
HSERO_RS24130	complement (2278837..2279160)
HSERO_RS10095	2281088..2281312
HSERO_RS10520	complement (2365679..2365879)
HSERO_RS10685	2404986..2406029
HSERO_RS10765	2420397..2420678
HSERO_RS10955	2469888..2470373
HSERO_RS11245	2539760..2540044
HSERO_RS11295	2553141..2553485
HSERO_RS11310	2556286..2556522
HSERO_RS11325	2558014..2558586
HSERO_RS11330	2558707..2558949
HSERO_RS11340	complement (2559970..2560194)
HSERO_RS11370	2567559..2568887
HSERO_RS11415	2579393..2579599
HSERO_RS11475	2592123..2592302
HSERO_RS11600	complement (2617572..2617778)
HSERO_RS11785	2692870..2693079
HSERO_RS11790	2693381..2693590
HSERO_RS11825	2698681..2698947
HSERO_RS11830	2699196..2699588
HSERO_RS11880	2707267..2707449
HSERO_RS11920	complement (2718649..2718924)
HSERO_RS12145	complement (2759300..2759617)
HSERO_RS12155	complement (2761139..2761348)
HSERO_RS12195	2769540..2769728

Locus_tag_HGF	Localização
SC_1_N_306	complement (3997046..3997300)
SC_1_N_307	complement (3997673..3997966)
SC_1_N_308	4046496..4046660
SC_1_N_309	complement (4059457..4059624)
SC_1_N_310	complement (4059919..4060041)
SC_1_N_311	4062693..4062998
SC_1_N_312	complement (4075097..4075267)
SC_1_N_316	4116904..4117302
SC_1_N_317	4126398..4126490
SC_1_N_318	4126535..4126684
SC_1_N_320	complement (4148686..4148829)
SC_1_N_321	complement (4152136..4152276)
SC_1_N_326	complement (4235570..4235659)
SC_1_N_333	complement (4316694..4316783)
SC_1_N_341	4451537..4451692
SC_1_N_342	4471054..4471131
SC_1_N_343	4471165..4471302
SC_1_N_350	complement (4553997..4554179)
SC_1_N_364	complement (4759196..4759306)
SC_1_N_389	complement (5118582..5118734)
SC_1_N_391	5188949..5189107
SC_1_N_394	5225492..5225746
SC_1_N_395	5239982..5240134
SC_1_N_396	complement (5240216..5240311)
SC_1_N_407	5335763..5335939
SC_1_N_415	complement (5399225..5399383)
SC_1_N_417	complement (5439484..5439570)
SC_1_N_419	5485201..5485329

Locus_tag_NCBI	Posição
HSERO_RS12260	complement (2783538..2784146)
HSERO_RS12455	complement (2833227..2833601)
HSERO_RS12525	2857713..2857895
HSERO_RS12640	complement (2883247..2883684)
HSERO_RS12900	complement (2940334..2940606)
HSERO_RS12940	2950884..2952431
HSERO_RS12950	2953401..2954933
HSERO_RS13210	complement (3011857..3012084)
HSERO_RS13225	3015403..3015600
HSERO_RS13315	complement (3037149..3037490)
HSERO_RS13330	complement (3039036..3039572)
HSERO_RS13340	complement (3040669..3040851)
HSERO_RS13345	complement (3040859..3041140)
HSERO_RS13365	complement (3045357..3045608)
HSERO_RS13370	3046050..3046430
HSERO_RS13425	complement (3056260..3056502)
HSERO_RS13430	complement (3056551..3056799)

Locus_tag_HGF	Localização
SC_1_N_420	complement (5492864..5493106)
SC_1_N_422	complement (5511706..5511984)
Locus_tag_Glimmer	**Localização**
ADJ62412.1	complement (971700..971846)
ADJ62508.1	complement (1082642..1082797)
ADJ62564.1	1142356..1142508
ADJ62738.1	1357294..1357434
ADJ62804.1	complement (1425741..1425854)
ADJ62941.1	1590856..1590996
ADJ63099.1	1773502..1773669
ADJ63428.1	2123191..2123328
ADJ63511.1	2224072..2224212
ADJ63815.1	complement (2547431..2547814)
ADJ64392.1	complement (3202577..3202690)
ADJ66206.1	complement (5303761..5304093)

Locus_tag_NCBI	Posição
HSERO_RS13460	complement (3061181..3061495)
HSERO_RS13830	3147032..3147283
HSERO_RS13880	complement (3156480..3156659)
HSERO_RS13885	3156852..3157514
HSERO_RS13985	3183458..3183541
HSERO_RS24160	complement (3264380..3267151)
HSERO_RS14425	3267166..3267726
HSERO_RS14430	complement (3267797..3269974)
HSERO_RS24165	3269988..3270659
HSERO_RS14440	complement (3270661..3271026)
HSERO_RS14605	3307064..3307255
HSERO_RS14790	complement (3357998..3358642)
HSERO_RS14860	complement (3374455..3374661)
HSERO_RS14910	complement (3387406..3387771)
HSERO_RS15045	3415743..3416012
HSERO_RS15050	complement (3416165..3416425)
HSERO_RS15620	3568644..3569210
HSERO_RS15635	complement (3571279..3571539)
HSERO_RS15720	complement (3593103..3593336)
HSERO_RS15855	complement (3621558..3621818)
HSERO_RS15920	complement (3633428..3633772)
HSERO_RS16045	complement (3663743..3663937)
HSERO_RS16080	3671609..3672160
HSERO_RS16250	complement (3705635..3705940)
HSERO_RS16380	3751157..3751744
HSERO_RS16510	complement (3779873..3780004)

Locus_tag_NCBI	Posição
HSERO_RS16565	3791251..3791679
HSERO_RS16595	complement (3800185..3800415)
HSERO_RS16635	complement (3809456..3809725)
HSERO_RS16865	3859311..3859607
HSERO_RS16910	3869895..3870359
HSERO_RS16940	complement (3877433..3877651)
HSERO_RS17095	complement (3915700..3915954)
HSERO_RS18045	complement (4148461..4148676)
HSERO_RS18080	complement (4155958..4156191)
HSERO_RS18100	4159098..4159682
HSERO_RS18105	4159994..4160560
HSERO_RS18535	4253215..4253688
HSERO_RS18610	4268960..4269334
HSERO_RS18630	4274789..4275190
HSERO_RS18695	4291735..4292130
HSERO_RS18740	4301721..4302164
HSERO_RS18835	4321295..4321834
HSERO_RS18845	4322879..4323451
HSERO_RS18875	complement (4331877..4331966)
HSERO_RS19030	complement (4363075..4364154)
HSERO_RS19305	complement (4428528..4428992)
HSERO_RS19575	4485213..4485821
HSERO_RS19705	complement (4522321..4522644)
HSERO_RS19750	complement (4530542..4530853)
HSERO_RS19770	complement (4535378..4535647)
HSERO_RS20080	complement (4607552..4607785)

Locus_tag_NCBI	Posição
HSERO_RS20125	4618814..4619245
HSERO_RS20210	4640811..4641116
HSERO_RS20360	complement (4665326..4665562)
HSERO_RS20480	4694881..4695195
HSERO_RS20500	4698079..4698540
HSERO_RS20565	4715676..4717007
HSERO_RS20800	4762834..4763061
HSERO_RS20840	4769752..4770139
HSERO_RS21245	4864576..4865232
HSERO_RS21630	4952646..4952855
HSERO_RS21735	complement (4976131..4976415)
HSERO_RS21990	complement (5030764..5030958)
HSERO_RS22885	complement (5252533..5252757)
HSERO_RS22920	complement (5259641..5259928)
HSERO_RS22935	complement (5264092..5264271)
HSERO_RS23120	complement (5308167..5308577)
HSERO_RS23220	5328032..5328229
HSERO_RS23480	complement (5385613..5385825)
HSERO_RS23490	5387434..5387703
HSERO_RS24030	5511315..5511548

Anexo 06: Operons preditos pelo *software* Rockhopper ou pelo Banco de Dados DOOR e corrigidos após reanotação.

Inicio	**Fim**	***Strand***	**Nº de genes**	**Genes**
23342	25449	+	2	czcR, czcS
29277	34816	+	3	Hsero_0020, Hsero_0021, Hsero_0022
34891	37553	+	2	Hsero_0023, Hsero_0024
37678	39643	+	3	Hsero_0025, Hsero_0026, Hsero_0027
42910	44049	-	2	Hsero_0031, Hsero_0032
46587	48159	-	2	fecR, fecI
63766	66169	+	3	Hsero_0049, Hsero_0050, feoB
69098	71909	-	3	tauC, tauB, tauA
74416	76379	+	2	Hsero_0061, speB
77496	78976	+	2	glnP, glnQ
79037	80577	-	2	Hsero_0066, Hsero_0067
82258	86458	-	3	ntrY, Hsero_0071, rrmB
89932	95807	-	6	cdsA, plsC, HSERO_RS00385, Hsero_0078, Hsero_0079, Hsero_0080
98251	100938	+	3	Hsero_0083, glnK, amtB
101286	104402	+	4	gshA, gshB, pts, ptsH
106756	108494	+	2	metX, Hsero_0092
111797	112769	+	2	secE, nusG
112930	114056	+	2	rplK, rplA
124609	128080	+	2	cirA, hemS
128410	128793	+		rpsL, rpsG, fusA, tufB
133522	138307	+	10	rplC, rplD, rplW, rplB, rpsS, rplV, rpsC, rplP, rpmC, rpsQ
138504	145153	+	15	rplN, rplX, rplE, rpsN, rpsH, rplF, rplR, rpsE, rpmD, rplO, secY, infA, rpmJ, rpsM, rpsK
147754	149942	+	2	cutA, dsbD
156854	158874	-	2	rpoE, hemB
166667	168133	-	2	gpsA, Hsero_0159
168798	169468	-	2	grxC, Hsero_0162
170458	173370	+	2	yibP, prc
181936	183900	+	2	Hsero_0175, Hsero_0176
185026	187023	+	2	livM, livG
192094	194267	+	3	Hsero_0182, Hsero_0183, Hsero_0184
200180	201089	-	2	gspG, cyaY
206791	209972	+	2	Hsero_0195, otsA
210152	211162	+	2	Hsero_0197, Hsero_0198
213343	215382	-	2	Hsero_0200, Hsero_0201
215425	219103	-	5	Hsero_0202, Hsero_0203, Hsero_0204, Hsero_0205, Hsero_0206
220331	223091	-	4	Hsero_0209, Hsero_0210, Hsero_0211, Hsero_0212

Inicio	Fim	*Strand*	Nº de genes	Genes
223160	227973	-	5	Hsero_0213, Hsero_0214, Hsero_0215, Hsero_0216, Hsero_0217
228196	230679	-	4	HSERO_RS01075, Hsero_0219, Hsero_0220, Hsero_0221
231806	236407	+	5	Hsero_0223, Hsero_0224, Hsero_0225, Hsero_0226, exeA, Hsero_0228
236489	238032	+	4	Hsero_0229, Hsero_0230, Hsero_0231, Hsero_0232
238395	240934	+	2	Hsero_0233, Hsero_0234
241053	244530	+	4	Hsero_0235, Hsero_0236, Hsero_0237, Hsero_0238
245979	250174	-	5	Hsero_0240, Hsero_0241, Hsero_0242, Hsero_0243, lldD
252738	254993	-	2	Hsero_0247, Hsero_0248
255099	259957	-	2	Hsero_0249, Hsero_0250
261271	263781	+	2	Hsero_0253, Hsero_0254
276107	279383	+	3	pta, ackA, phaC
280207	285958	+	3	Hsero_0267, Hsero_0268, Hsero_0269
286019	287832	-	2	Hsero_0270, Hsero_0271
296150	302528	+	3	Hsero_0279, Hsero_0280, Hsero_0281
289392	292609	-	2	lacA, Hsero_0274
302588	304238	+	2	Hsero_0282, Hsero_0283
308723	313118	-	3	Hsero_0289, Hsero_0290, Hsero_0291
313202	314319	-	2	Hsero_0292, exbD
325445	327225	+	2	tktA1, tktA2
328779	335539	+	3	Hsero_0306, Hsero_0307, Hsero_0308
342527	349598	-	4	Hsero_0314, Hsero_0315, Hsero_0316, Hsero_0317
359967	376572	+	14	mraZ, mraW, ftsL, ftsI, murE, murF, mraY, murD, ftsW, murG, murC, ddlB, ftsQ, ftsA
384760	385601	+	2	Hsero_0347, Hsero_0348
386389	388951	+	3	argJ, Hsero_0350, mutT
388988	389958	-	2	Hsero_0352, Hsero_0353
390111	394025	-	4	coaE, pilD, pilC, pilB
394189	395748	-	2	corB, Hsero_0359
395922	396546	+	2	infA, Hsero_0361
398080	398697	+	2	rplU, rpmA
401333	402501	-	2	Hsero_0368, rppH
407919	410530	+	3	Hsero_0374, Hsero_0375, Hsero_0376
410664	412384	+	3	Hsero_0377, Hsero_0378, ampD
421059	423266	-	2	Hsero_0386, prmA
423308	425256	-	2	accC, accB
425925	427087	-	2	tlpA, Hsero_0392
428446	436150	+	7	cheD1, Hsero_0395, mpl, Hsero_0397, ubiC, vacB, tonB
436239	437788	+	2	aroE, mtgA
442429	444276	+	2	tyrS, dtd
444363	445664	+	2	Hsero_0409, gpmB
445806	447436	+	2	hipB, hipA

Inicio	Fim	*Strand*	Nº de genes	Genes
452144	454479	+	3	phrB, Hsero_0418, ruvC
455531	457204	+	2	ruvA, ruvB
457384	458701	+	2	Hsero_0423, Hsero_0424
462763	467233	+	5	thiC, HSERO_RS02125 (thiS), thiG, thiD, thiE
475745	477672	-	2	pfkB, dalR
479558	483243	+	4	Hsero_0441, Hsero_0442, Hsero_0443, Hsero_0444
484958	486911	+	2	xylB, Hsero_0447
488946	490717	-	2	Hsero_0449, thrB
499593	502738	+	3	Hsero_0457, Hsero_0458, pcbC
502854	506991	+	4	hcaE, tauB, tauC, ssnA
507263	509706	+	4	Hsero_0464, Hsero_0465, pyrR, pyrB
509882	511452	+	2	Hsero_0468, pyrX
513348	515747	+	2	cpsG, rfaC
515907	518062	+	2	Hsero_0474, kdtA
525480	531300	-	3	Hsero_0482, acrB, acrA
533252	534260	-	2	Hsero_0487, pcm
541924	545370	-	4	Hsero_0498, Hsero_0499, Hsero_0500, Hsero_0501
545545	547139	+	2	HSERO_RS02490, Hsero_0502, Hsero_0503
554479	556505	-	2	Hsero_0508, Hsero_0509
563807	565854	+	2	Hsero_0516, Hsero_0517
566985	568061	-	2	Hsero_0519, Hsero_0520
568141	569331	-	2	Hsero_0521, Hsero_0522
577257	579294	-	2	Hsero_0529, Hsero_0530
587597	589013	-	2	gloA, Hsero_0536
592577	595680	-	2	Hsero_0539, Hsero_0540
597730	600487	-	2	Hsero_0543, Hsero_0544
600817	602435	+	2	Hsero_0545, Hsero_0546
615746	617784	+	2	Hsero_0559, metE
619344	621131	+	2	Hsero_0563, Hsero_0564
623449	629906	-	4	Hsero_0568 (ndeD), Hsero_0569, ndeD, Hsero_0571
632110	633146	+	2	Hsero_0574, Hsero_0575
634325	635916	-	2	Hsero_0577, Hsero_0578
640650	642759	-	2	Hsero_0582, Hsero_0583
643002	644275	-	2	fabG, Hsero_0585
644592	646382	-	2	Hsero_0586, Hsero_0587
648436	655159	-	6	nahD, Hsero_0591, proA, holA, rlpB, leuS
644592	661383	+	7	Hsero_0596, Hsero_0597, Hsero_0598, Hsero_0599, potC, potA, Hsero_0602
661659	663164	-	3	Hsero_0603, exbD, exbB
663324	664689	-	2	dapB, smpA
668280	671080	+	2	recN, hemH
675786	677262	-	2	panB, Hsero_0618
677904	679187	+	2	Hsero_0619, cheY
679305	687471	+	7	Hsero_0621, Hsero_0622, Hsero_0623, cheW, cheR,

Inicio	Fim	*Strand*	Nº de genes	Genes
				cheB, Hsero_0627
687520	691013	-	4	folK, pcnB, serB, hda
694571	697001	-	2	Hsero_0635, Hsero_0636
700133	707060	+	3	Hsero_0640, Hsero_0641, Hsero_0642
707081	710580	-	3	Hsero_0643, Hsero_0644, Hsero_0645
714050	716961	-	3	Hsero_0649, Hsero_0650, Hsero_0651
717026	719556	-	3	Hsero_0652, Hsero_0653, Hsero_0654
722392	724535	+	2	sucC, sucD
728422	730015	-	2	Hsero_0663, Hsero_0664
731951	733056	+	2	azlC, Hsero_0669
734597	737121	+	2	pykA, fbaA
737213	740872	+	4	purC, purE, purK, sua5
745598	747504	-	2	cfa, pdxH
752891	755381	-	5	exbD, exbB, tonB, bfd, Hsero_0693
755883	757526	-	2	trmU, Hsero_0695
757691	761075	+	4	pntAa, panD, pntAb, pntB
763194	766295	-	3	pepP, Hsero_0703, Hsero_0704
766393	770207	+	2	ostA, surA
770303	772134	+	2	pdxA, ksgA
778121	779541	-	2	plsC, Hsero_0715
782859	785386	-	2	lnt, corC
779632	782747	-	2	glyS, glyQ
787626	790564	-	3	Hsero_0721, phoH, miaB
792807	798348	+	5	Hsero_0726, Hsero_0727, Hsero_0728, Hsero_0729, Hsero_0730
798392	800377	-	2	Hsero_0731, Hsero_0732
800673	802962	-	2	Hsero_0733, Hsero_0734, Hsero_0735
803061	808097	-	3	Hsero_0736, Hsero_0737, Hsero_0738
809942	812465	+	3	Hsero_0741, Hsero_0742, Hsero_0743
812608	818708	+	4	Hsero_0744, Hsero_0745, Hsero_0746, clpB
821697	825866	+	2	Hsero_0749, Hsero_0750
827071	829105	+	2	Hsero_0752, Hsero_0753
831617	839148	-	5	Hsero_0756, Hsero_0757, Hsero_0758, Hsero_0759, Hsero_0760
840613	849568	+	8	HSERO_RS03805, HSERO_RS03810, Hsero_0763, Hsero_0764, Hsero_0765, Hsero_0766, Hsero_0767, Hsero_0768
862350	867028	+	6	Hsero_0781, Hsero_0782, Hsero_0783, Hsero_0784, Hsero_0785, hrcV
867213	868645	+	2	Hsero_0787, Hsero_0788
868739	874196	+	6	hrpQ, hrcN, Hsero_0791, Hsero_0792, hrcQ, hrcR
874262	876981	+	4	hrcS, hrpX, hrcU, Hsero_0798
877060	879002	+	3	hrpB, hrcJ, hrpD
879871	882600	+	3	Hsero_0803, Hsero_0804, hrcC
883186	890145	+	6	Hsero_0806, pilN, pilO, pilP, gspE, bfpE

Inicio	Fim	*Strand*	Nº de genes	Genes
890862	893237	+	3	Hsero_0813, pilV, Hsero_0815
894726	897193	+	3	Hsero_0817, Hsero_0818, Hsero_0819
904585	907129	+	3	Hsero_0825 Hsero_0826, Hsero_0827
910154	915144	-	2	Hsero_0830, Hsero_0831
915600	918431	-	2	miaA, mutL
925067	926877	-	2	amiC, Hsero_0839
928103	931677	-	3	Hsero_0841, dctR, dctS
933287	935124	+	2	dctQ, dctM
936348	937817	+	2	ogT, xerD
937896	939045	+	2	Hsero_0850, Hsero_0851
951188	957435	+	2	Hsero_0861, uvrD
962832	967692	+	3	dnaG, Hsero_0869, rpoD
974612	975295	-	3	Hsero_0879, Hsero_0880, Hsero_0881
976834	981228	-	3	Hsero_0884, Hsero_0885, Hsero_0886
981734	989295	-	10	Hsero_0888, Hsero_0889, Hsero_0890, Hsero_0891, Hsero_0892, Hsero_0893, Hsero_0894, Hsero_0895, Hsero_0896, Hsero_0897
981734	995632	-	11	Hsero_0898, Hsero_0899, Hsero_0900, Hsero_0901, Hsero_0902, Hsero_0903, Hsero_0904, Hsero_0905, Hsero_0906, Hsero_0907, Hsero_0908
992978	994740	-	2	
995776	998824	+	3	Hsero_0909, Hsero_0910, Hsero_0911
999952	1002531	-	5	Hsero_0914, Hsero_0915, Hsero_0916, Hsero_0917, Hsero_0918
1002956	1004715	+	2	Hsero_0920, Hsero_0921
1004985	1006417	-	2	Hsero_0922, Hsero_0923
1006546	1011495	+	4	acs, Hsero_0925, atoB, Hsero_0927
1012983	1016616	-	4	oxdA, Hsero_0930, nthB, nthA
1028934	1034361	+	4	Hsero_0941, Hsero_0942, Hsero_0943, Hsero_0944
1042184	1046797	+	5	Hsero_0950, Hsero_0951, livG, livH, livM
1048249	1050271	+	2	Hsero_0956, livF
1051448	1053844	+	2	mmsB, Hsero_0960
1059643	1065407	+	6	glpD2, Hsero_0967, Hsero_0968, Hsero_0969, Hsero_0970, Hsero_0971
1084227	1088762	-	2	Hsero_0985, fbpC, mltB
1094540	1096220	-	2	Hsero_0993, Hsero_0994
1101299	1103180	+	2	mdpB, Hsero_0999
1103337	1109749	+	2	fhaB, fhaC
1111065	1115674	+	5	Hsero_1003, Hsero_1004, Hsero_1005, Hsero_1006, Hsero_1007
1115752	1120278	+	5	Hsero_1008, Hsero_1009, Hsero_1010, Hsero_1011, Hsero_1012
1121086	1122436	+	2	Hsero_1013, Hsero_1014
1123508	1129717	-	6	Hsero_1016, Hsero_1017, Hsero_1018, Hsero_1019, rbsK, rbsD
1129918	1133357	-	2	Hsero_1022, deoC

Inicio	Fim	*Strand*	Nº de genes	Genes
1139550	1142168	-	3	dgoAa, dgoK, galM
1143757	1147402	+	3	Hsero_1034, Hsero_1035, Hsero_1036
1148902	1154450	+	4	Hsero_1038, Hsero_1039, Hsero_1040, Hsero_1041
1160755	1161901	+	2	Hsero_1047, Hsero_1048
1162060	1165728	+	3	Hsero_1049, Hsero_1050, Hsero_1051
1167539	1172721	+	6	Hsero_1054, Hsero_1055, cfa, Hsero_1057, Hsero_1058, Hsero_1059
1174765	1178224	+	3	Hsero_1062, Hsero_1063, rbsC
1178277	1181436	+	3	rbtD, araB, Hsero_1067
1182029	1183845	+	2	metX, Hsero_1070
1186452	1189045	+	3	Gst, Hsero_1074, Hsero_1075
1189154	1189779	+	2	HSERO_RS05385, Hsero_1076
1189835	1192202	-	3	gabD1, potC, potH
1194994	1197489	-	2	potA, goaG
1240197	1242057	+	2	idnO, idnD
1245187	1250329	-	4	talB, Hsero_1097, zwf, pgi
1250683	1253180	+	2	edd, eda
1255837	1257683	-	2	tktC, tktB
1257724	1260623	-	2	dctQM, dctP
1262870	1265257	-	2	Hsero_1111, Hsero_1112
1268220	1269997	-	2	Hsero_1114, Hsero_1115
1270200	1274903	+	4	pcaK, hmgA, mhpD, gst
1284135	1285969	-	2	serA, Hsero_1128
1289247	1293023	+	3	Hsero_1131, Hsero_1132, Hsero_1133
1293254	1295315	+	2	Hsero_1134, dadA
1296866	1299575	+	2	betB, hipO
1302734	1305409	-	2	Hsero_1142, Hsero_1143
1309872	1312207	-	2	Hsero_1147, Hsero_1148
1312428	1315553	-	2	Vdh, Hsero_1150
1316809	1317246	+	2	Hsero_1152, Hsero_1153
1318852	1321339	+	2	Hsero_1154, Hsero_1155
1327900	1330599	+	3	Hsero_1162, Hsero_1163, dadA
1333170	1334219	+	2	exbB2, exbD2
1339884	1341908	+	2	Hsero_1173, Hsero_1174
1342022	1342938	+	2	Hsero_1175, Hsero_1176
1343062	1345565	+	3	purN, sun, Hsero_1179
1348557	1348985	-	2	rpmG, rpmB
1352127	1355848	+	4	livH, livM, livG, livF
1357917	1358759	+	2	Bfr, Hsero_1196
1366062	1369885	+	4	nasF, nasE, nasD, Hsero_1210
1374750	1379488	-	6	ybgF, HSERO_RS06105, Hsero_1217, tolA, exbD, tolQ
1388011	1390187	+	2	tsaA, tgt
1390210	1392643	-	2	vsrB, corR
1392802	1396970	-	3	Hsero_1232, Hsero_1233, Hsero_1234

Inicio	Fim	*Strand*	Nº de genes	Genes
1399618	1402550	+	2	secD, secF
1413166	1415940	+	3	Hsero_1247, Hsero_1248, Hsero_1249
1422925	1425563	-	2	sixA, ppk
1427253	1429926	+	3	pstC, pstA, pstB
1430101	1432995	+	3	phoU, phoB, phoR
1442586	1444078	+	2	Hsero_1271, nahD
1444068	1445404	-	3	Hsero_1273, Hsero_1274, Hsero_1275
1450118	1450393	+	4	HSERO_RS06420, maoN, Hsero_1280, Hsero_1281
1458931	1460014	+	2	Hsero_1286, Hsero_1287
1461855	1462293	+	2	Hsero_1289, Hsero_1290
1462341	1462592	+	2	Hsero_1291, Hsero_1292
1463867	1474428	+	3	hlyB, Hsero_1294, Hsero_1295
1474494	1478437	+	2	Hsero_1296, Hsero_1297
1479980	1483545	+	2	Hsero_1300, Hsero_1301
1485791	1486874	-	2	catD, catC
1488133	1494592	+	7	catB2, isoA, cbeA, cbeB, cbeC, cbeD, yjgH
1496382	1500394	+	3	benE, Hsero_1315, Hsero_1316
1501173	1503204	+	3	Hsero_1318, Hsero_1319, mopI
1503226	1506629	-	2	Hsero_1321, iorA
1507031	1512733	+	5	Hsero_1323, Hsero_1324, Hsero_1325, entB, coxL
1516212	1518320	+	2	Hsero_1331, xdhC
1526061	1530379	-	5	Hsero_1339, hcaA2, hcaA1, Hsero_1342, HSERO_RS06730
1534567	1538169	+	4	ncnH, Hsero_1347, fabG, Hsero_1349
1538307	1543856	+	5	Hsero_1350, Hsero_1351, Hsero_1352, Hsero_1353, Hsero_1354
1545436	1551916	+	4	Hsero_1357, Hsero_1358, tolC, Hsero_1360
1555348	1558933	+	2	Hsero_1363, Hsero_1364
1559281	1562847	-	2	Hsero_1365, Hsero_1366
1567754	1569393	+	2	Hsero_1372, N_115
1569333	1570901	-	2	Hsero_1373, Hsero_1374
1573710	1578099	+	2	carA, carB
1582725	1585759	+	3	folP, mrsA, phoB
1585859	1587909	-	2	cccA, Hsero_1388
1600152	1602766	+	2	clpS, clpA
1602936	1604494	-	2	Hsero_1400, dut
1604568	1610381	-	4	coaBC, lspA, ileS, ribF
1610570	1615577	-	6	Hsero_1406, Hsero_1407, Hsero_1408, Hsero_1409, Hsero_1410, Hsero_1411
1621064	1623600	+	2	tktA, Hsero_1416
1634038	1636330	+	2	Hsero_1422, tldD
1639760	1642433	-	3	Hsero_1427, argG, argF
1644953	1647595	-	3	Hsero_1431, Hsero_1432, Hsero_1433
1649226	1650827	+	2	thrB, tag
1666013	1666903	-	2	Hsero_1448, holC

Inicio	Fim	*Strand*	Nº de genes	Genes
1667247	1670237	-	2	prlC, folD
1675515	1679000	+	2	aceF, lpdA
1681730	1683855	-	2	Hsero_1458, Hsero_1459
1685176	1686594	-	3	Hsero_1461, Hsero_1462, Hsero_1463
1689823	1691654	-	3	Hsero_1467, Hsero_1468, ttuC
1698724	1704604	+	4	fdsG, fdsB, fdhF, fdsC
1709705	1712280	-	3	tauC, tauB, tauA
1714024	1715129	-	3	Hsero_1490, Hsero_1491, Hsero_1492
1720081	1723840	+	5	Hsero_1496, Hsero_1497, Hsero_1498, Hsero_1499, Hsero_1500
1729579	1729833	+	2	rpsP, rimM
1725186	1726873	+	2	Hsero_1502, etfA
1738777	1740917	-	2	envZ, ompR
1742253	1751608	+	4	Hsero_1520, Hsero_1521, Hsero_1522, Hsero_1523
1753733	1757016	+	3	Hsero_1525, Hsero_1526, corA
1757275	1764299	+	5	cirA, Hsero_1529, Hsero_1530, Hsero_1531, Hsero_1532
1764311	1766335	-	2	Hsero_1533, Hsero_1534
1776071	1779223	-	3	murI, fumA, Hsero_1544
1782149	1783903	+	3	Hsero_1547, Hsero_1548, Hsero_1549
1785379	1787634	+	2	Hsero_1552, Hsero_1553
1787738	1788385	-	2	Hsero_1554, cheY
1793752	1795448	-	2	Hsero_1559, Hsero_1560
1798181	1803590	+	5	Hsero_1563, dppB2, dppC2, dppD2, dppF2
1804051	1808542	+	5	ddpA, Hsero_1569, Hsero_1570, Hsero_1571, gcvH
1811697	1813122	+	2	Hsero_1578, Hsero_1579
1814097	1814402	+	2	Hsero_1581, Hsero_1582
1816826	1817994	+	2	Hsero_1585, ssuA
1821726	1823433	+	3	Hsero_1591, Hsero_1592, purU
1823426	1823766	-	2	Hsero_1594, Hsero_1595
1826184	1832219	-	5	Hsero_1600, copA, copB, Hsero_1603, zntA
1833793	1836457	-	2	soxR, copA
1836951	1839626	-	3	Hsero_1610, rimL, moxR
1839846	1842274	+	2	Hsero_1614, Hsero_1615
1842429	1843117	+	2	Hsero_1616, Hsero_1617
1845024	1846393	-	2	nth, fdx
1842830	1855313	+	2	arsB, Hsero_1631
1846523	1847921	-	2	phaZ, Hsero_1623
1855428	1856060	-	2	acb, fkpA
1856370	1858206	+	2	Hsero_1634, Hsero_1635
1858303	1860131	+	2	Hsero_1636, Hsero_1637
1862054	1863945	+	2	Hsero_1640, Hsero_1641
1868690	1870361	+	2	sbp, Hsero_1647
1870471	1874353	+	4	ssuA, ssuD, ssuC, ssuB

Inicio	Fim	*Strand*	Nº de genes	Genes
1874694	1876139	+	2	Hsero_1653, Hsero_1654
1876303	1880115	+	4	cysU, cysW, cysA, cysB
1883635	1888537	+	5	livK, livH, livM, livG, livF
1888765	1892607	+	3	Hsero_1667, Hsero_1668, livF
1892829	1896309	+	3	Hsero_1670, Hsero_1671, Hsero_1672
1901916	1903474	+	2	Hsero_1679, alkB
1909337	1911961	+	2	Hsero_1686, HSERO_RS0390
1922605	1935895	+	11	Hsero_1698, cheW, cheB, Hsero_1701, cheR, cheY, cheA, HSERO_RS08485, cheY, Hsero_1706, Hsero_1707
1936203	1939572	+	2	Hsero_1708, Hsero_1709
1943838	1947338	+	4	opuBB1, opuBC, opuBB2, opuA
1949471	1951549	-	2	Hsero_1719, rhlE
1951717	1953020	-	2	Hsero_1721, Hsero_1722
1954230	1955781	+	2	lasT, rsuA
1955902	1956306	-	2	N_142, HSERO_RS08600
1956412	1957886	-	2	Hsero_1726, Hsero_1727
1959449	1960980	-	2	Hsero_1730, Hsero_1731
1961222	1963090	-	3	Hsero_1732, Hsero_1733, fecI
1963637	1966941	+	3	ilvI, ilvH, ilvC
1967166	1968606	+	2	Hsero_1738, Hsero_1739
1968784	1970372	+	2	psd, pssA
1971086	1973310	-	2	Hsero_1743, Hsero_1744
1977113	1979351	-	2	Hsero_1747, dsd
1979520	1981896	+	2	Hsero_1749, ndeD
1989168	1991952	-	3	Hsero_1757, Hsero_1758, Hsero_1759
1992222	1994653	-	3	Hsero_1760, Hsero_1761, Hsero_1762
1994910	1996436	+	2	Hsero_1763, qor
1998086	2000784	+	4	nuoA, nuoB, nuoC, nuoD
2000892	2006647	+	5	nuoE, nuoF, nuoG, nuoH, nuoI
2006788	2007774	+	2	nuoJ, nuoK
2007890	2011493	+	2	nuoL, nuoM
2011652	2013487	+	2	nuoN, Hsero_1781
2018270	2021863	+	3	livH, livG, livF
2028992	2030084	-	2	Hsero_1799, Hsero_1800
2032590	2036388	+	3	Hsero_1802, Hsero_1803, Hsero_1804
2036513	2039914	+	2	Hsero_1805, Hsero_1806
2049691	2052304	+	3	ilvA, rhtB, Hsero_1811
2052477	2054573	+	2	tdcF, Hsero_1813, Hsero_1814
2054690	2055458	-	2	Hsero_1815, Hsero_1816
2058296	2059954	-	3	Hsero_1820, Hsero_1821, Hsero_1822
2071679	2078355	+	5	Hsero_1829, Hsero_1830, yhiH, yhhJ, N_152
2078428	2081531	-	2	parC, Hsero_1834
2089197	2093042	+	3	Hsero_1840, Hsero_1841, Hsero_1842

Inicio	Fim	*Strand*	Nº de genes	Genes
2094884	2098642	+	4	dppB, dppC, dmpA, dppA
2098744	2099678	+	2	Hsero_1848, Hsero_1849
2111696	2115063	+	2	vacB, sopU
2104379	2105603	-	2	mog, Hsero_1856
2115996	2117306	-	2	Hsero_1865, ppiB2
2117436	2119201	-	2	ppiB1, Hsero_1868
2121108	2123160	+	2	cysS, alkA
2126120	2126813	-	2	Hsero_1874, Hsero_1975
2134055	2137159	-	3	moeA, mobA, moaA
2142567	2143610	+	2	gph, Hsero_1888
2179539	2181640	+	2	plsX, fabH
2181714	2183398	+	2	fabD, fabG
2183562	2185076	+	2	acpP, fabF1
2186432	2189641	+	3	rseA, rseB, degQ
2190206	2192915	+	2	lepA, lepB
2193109	2194994	+	2	rnc, era
2196105	2200192	+	4	pdxJ, acpS, Hsero_1941, uvrC
2201625	2208021	-	4	Hsero_1948, Hsero_1949, Hsero_1950, tolC
2208252	2210158	-	2	Hsero_1952, Hsero_1953
2210391	2214036	+	3	Hsero_1954, Hsero_1955, Hsero_1956
2215247	2217749	+	3	cobT, cobS, gpmB
2217756	2219808	-	2	Hsero_1961, Hsero_1962
2220043	2222347	+	2	cstA, Hsero_1964
2229810	2232416	+	2	yjgF, relA
2237074	2237659	+	2	rpmI, rplT
2237761	2241385	+	2	pheS, pheT
2241549	2242274	+	2	ihfA, Hsero_1984
2250071	2252618	+	2	epsG, epsH
2252662	2255381	+	2	epsM, Hsero_1995
2265440	2269117	+	3	ugd, Hsero_2004, wzxC
2269173	2271019	+	2	epsJ, Hsero_2007
2272287	2273881	+	2	cysE, galE
2281390	2282666	-	2	czcD, Hsero_2018
2284732	2287346	+	3	cheR, cheD, cheB
2287520	2288567	+	2	cheY, cheZ
2288842	2295193	+	5	flhB, flhA, flhF, fleN, fliA
2295283	2296131	-	2	flgN, Hsero_2031
2297196	2300106	+	4	flgB, flgC, flgD, flgE
2300221	2304740	+	5	flgF, flgG, flgH, flgI, flgJ
2304879	2308370	+	2	flgK, flgL
2308468	2313082	-	3	Hsero_2044, Hsero_2045, Hsero_2046
2313193	2315061	-	3	fliR, fliQ, fliP
2315108	2317755	-	4	fliO, fliN, fliM, fliL
2319235	2321859	-	3	fliJ, fliI, fliH

Inicio	Fim	*Strand*	Nº de genes	Genes
2321913	2324815	-	3	fliG, Hsero_2059, Hsero_2060
2325410	2327568	-	3	Hsero_2062, flhB, Hsero_2064
2327706	2330606	-	3	fliT, fliS, fliD
2332419	2334669	-	2	Hsero_2071, Hsero_2072
2335914	2336600	+	2	rpsF, priB
2336656	2337419	+	2	rpsR, rplI
2339124	2339901	+	2	Hsero_2079, Hsero_2080
2344929	2347884	-	3	Hsero_2085, fmt, pmrF
2349308	2351829	-	3	Hsero_2089, arnT, Hsero_2091
2351990	2353460	+	2	Hsero_2092, aspC
2357288	2359496	+	3	moeA, moaD, Hsero_2099
2361067	2362930	+	2	moaE, mntH
2363007	2364584	+	2	Hsero_2103, Hsero_2104
2365869	2369533	-	2	Hsero_2106, Hsero_2107
2369742	2371119	+	2	Hsero_2108, crcB
2374338	2376544	+	2	Hsero_2111, menG
2380684	2382843	-	3	modC1, modB1, modA1
2387815	2389312	-	2	Hsero_2122, mgsA
2393546	2396046	-	2	spoT, rpoZ
2398951	2400782	+	2	Hsero_2134, hemN
2400925	2404181	+	2	Hsero_2136, Hsero_2137
2410079	2413353	+	4	Hsero_2145, Hsero_2146, Hsero_2147, Hsero_2148
2414906	2417448	+	2	dadX, radA
2417464	2420177	-	3	Hsero_2152, risS, risA
2422494	2426934	-	4	Hsero_2158, hemK, dapE, arsC
2435464	2438879	+	2	Hsero_2166, lig
2440538	2445379	-	2	cphA2, cphA1
2445761	2448817	+	3	Hsero_2172, mdlB, Hsero_2174
2457269	2461432	+	4	uppS, cdsA, dxr, Hsero_2185
2461588	2463963	+	2	Hsero_2186, Hsero_2187
2464654	2466199	+	2	lpxD, fabZ
2466412	2468359	+	2	lpxA, lpxB
2468428	2469862	+	2	rnhB, spoU
2474597	2476870	+	2	Hsero_2197, Hsero_2198
2476968	2479080	-	2	Hsero_2199, Hsero_2200
2481554	2482283	+	2	Hsero_2203, Hsero_2204
2482414	2485589	+	2	guaB, guaA
2490362	2492366	+	3	Hsero_2211, Hsero_2212, ldcA
2493205	2496882	+	3	uup, nadA, nadC
2498263	2502148	-	2	Hsero_2219, nadB
2503899	2504566	+	2	Hsero_2222, Hsero_2223
2504591	2506250	-	2	Hsero_2224, Hsero_2225
2512412	2516357	-	3	Hsero_2230, Hsero_2231, Hsero_2232
2521991	2524457	-	2	Hsero_2238, Hsero_2239

Inicio	Fim	*Strand*	Nº de genes	Genes
2533050	2533858	-	2	Hsero_2244, ispZ
2536058	2540044	+	5	livH, livM, livG, livF, HSERO_RS11245
2547811	2552923	-	3	dnaE, Hsero_2259, Hsero_2260
2560353	2562389	-	2	Hsero_2268, poxB
2562602	2566218	+	2	Hsero_2270, Hsero_2271
2568938	2570769	-	2	Hsero_2274, Hsero_2275
2570825	2576774	-	4	Hsero_2276, Hsero_2277, Hsero_2278, Hsero_2279
2579718	2583526	-	2	Hsero_2282, acnA
2586315	2587753	+	2	glnP, glnQ
2593858	2598819	+	5	Hsero_2294, Hsero_2295, Hsero_2296, Hsero_2297, Hsero_2298
2609597	2615211	-	6	Hsero_2307, Hsero_2308, Hsero_2309, coxB, coxA, coxB
2615344	2617465	-	4	Hsero_2313, Hsero_2314, Hsero_2315, Hsero_2316
2618905	2623053	+	2	Hsero_2319, Hsero_2320
2624099	2626740	+	3	Hsero_2322, Hsero_2323, Hsero_2324
2627048	2636403	+	3	Hsero_2325, Hsero_2326, Hsero_2327
2637100	2638763	-	2	tauC, tauB
2640109	2642952	-	2	csdB, Hsero_2333
2643030	2644280	-	2	Hsero_2334, cysE
2645251	2648366	-	2	cirA, pfrI
2648532	2651225	+	4	Hsero_2339, Hsero_2340, Hsero_2341, Hsero_2342
2682281	2684638	-	2	Hsero_2346, Hsero_2347
2690246	2691813	+	2	Hsero_2353, Hsero_2354
2695932	2696646	+	2	Hsero_2361, Hsero_2362
2699585	2702021	+	5	Hsero_2367, Hsero_2368, Hsero_2369, Hsero_2370, Hsero_2371
2702453	2703092	-	2	Hsero_2372, Hsero_2373
2707572	2708902	+	2	Hsero_2377, Hsero_2378
2718734	2723431	+	5	Hsero_2385, Hsero_2386, Hsero_2387, Hsero_2388, Hsero_2389
2723460	2724171	-	2	Hsero_2390, Hsero_2391
2727385	2728318	+	2	Hsero_2397, Hsero_2398
2732028	2734898	+	3	dksA, phbC, fabI
2736489	2737701	-	2	Hsero_2408, Hsero_2409
2742290	2743293	+	2	Hsero_2414, Hsero_2415
2743328	2745238	-	2	Hsero_2416, Hsero_2417
2745286	2746518	-	2	Hsero_2418, rpoE
2747357	2749371	+	2	Hsero_2421, Hsero_2422
2751737	2759303	+	6	Hsero_2425, Hsero_2426, iolC, iolD, iolE, iolB
2761439	2763015	-	2	fecR, fecI
2763406	2765384	+	2	Hsero_2435, baeS
2778102	2784173	+	6	bioB, bioI, bioF, bioA, bioD, Hsero_2452
2786015	2789357	-	3	livF, livG, livH
2791393	2796730	-	5	hydH, shb1, cyoD, cyoC, cyoB

Inicio	Fim	*Strand*	Nº de genes	Genes
2805227	2807592	-	3	Hsero_2471, Hsero_2472, Hsero_2473
2811384	2814493	+	2	cls, Hsero_2477
2815243	2816670	-	2	Hsero_2479, lepB
2816891	2822850	+	3	yjcP, Hsero_2482, Hsero_2483
2822943	2825055	+	2	Hsero_2484, Hsero_2485
2825086	2826458	-	2	serA, Hsero_2487
2837739	2839849	-	2	Hsero_2494, Hsero_2495
2843031	2854770	-	5	treY, malQ, treZ, glgX, glgB
2859771	2861421	+	2	Hsero_2508, Hsero_2509
2864673	2865380	-	2	Hsero_2513, Hsero_2514
2868243	2870515	-	2	eutC, eutB
2873546	2875676	+	2	eutR, Hsero_2521
2879259	2880940	-	2	tauC, tauB
2881114	2883239	+	2	ygiX, ygiY
2888962	2891278	-	2	Hsero_2535, Hsero_2536
2892990	2895646	+	3	Hsero_2539, Hsero_2540, cpdA
2897704	2899832	+	2	Hsero_2544, Hsero_2545
2900896	2901842	-	2	Hsero_2547, Hsero_2548
2902001	2903962	-	2	caiD, caiA
2911275	2915956	+	3	Hsero_2555, Hsero_2556, gst
2919596	2921746	-	2	paaK, Hsero_2562
2925527	2928449	-	2	Hsero_2566, fadD
2930274	2934286	-	3	Hsero_2570, Hsero_2571, Hsero_2572
2935403	2938936	+	3	papD, papC, Hsero_2576
2945129	2947156	-	2	clpP, tig
2955951	2957079	+	2	Hsero_2589, crtB1
2958342	2960522	+	2	crtB, hpnE
2961977	2963424	-	2	Hsero_2595, Hsero_2596
2963484	2965399	-	2	Hsero_2597, Hsero_2598
2967838	2969494	-	2	holB, tmk
2971803	2972923	+	2	Hsero_2604, Hsero_2605
2974192	2978159	-	4	plcN, tonB, exbD, Hsero_2610
2981463	2985713	+	4	purR, ugpB, ugpA, ugpE
2991321	2992068	-	2	Hsero_2621, Hsero_2622
2996638	2997756	+	2	Hsero_2626, Hsero_2627
2999536	3004206	-	5	amiE, Hsero_2631, Hsero_2632, Hsero_2633, Hsero_2634
3006228	3010599	+	2	Hsero_2636, Hsero_2637
3015683	3025017	-	9	cobQ, cobC, cobD, cobP, nfnB, cobB, cobA, Hsero_2648, cbiJH
3025161	3037183	-	8	cbiG, cbiF, cbiL, cbiD, cobH, cobL, chlI, cobN
3037565	3038927	-	2	Hsero_2658, Hsero_2659
3056260	3056799	-	2	HSERO_RS13425, HSERO_RS13430
3056951	3061184	-	5	fabD, mdcB, mdcG, mdcC, accD

Inicio	Fim	*Strand*	Nº de genes	Genes
3061531	3065986	-	4	mdcA, dctM, dctQ, dctP
3066969	3069504	-	3	rluA, Hsero_2689, Hsero_2690
3070092	3070794	+	2	Hsero_2692, Hsero_2693
3071245	3072569	-	2	Hsero_2695, rpoE
3073140	3074611	-	2	Hsero_2698, Hsero_2699
3074677	3077715	-	2	Hsero_2700, Hsero_2701
3080483	3083321	+	2	emrA, amrB
3086406	3088050	-	3	Hsero_2709, Hsero_2710, Hsero_2711
3088310	3090275	-	2	Hsero_2712, adhC
3098376	3099713	-	2	napH, Hsero_2720
3107692	3116199	-	6	degQ, lysA, caiC, atoC, Hsero_2731, wcaJ
3117527	3119751	+	2	Hsero_2734, Hsero_2735
3119834	3123328	+	3	wecC, galE, Hsero_2738
3125270	3130839	+	5	Hsero_2740, Hsero_2741, Hsero_2742, wecB, Hsero_2744
3130944	3144115	+	10	Hsero_2745, Hsero_2746, Hsero_2747, Hsero_2748, asnB, Hsero_2750, Hsero_2751, Hsero_2752, Hsero_2753, Hsero_2754
3144249	3148972	+	4	Hsero_2755, HSERO_RS13830, Hsero_2756, Hsero_2757
3154113	3156454	-	2	Hsero_2763, Hsero_2764
3162662	3165183	-	2	kdsA, pyrG
3166289	3169584	-	2	Hsero_2774, Hsero_2775
3171182	3173150	-	2	lolD, lolE
3175950	3178672	+	2	Hsero_2780, recJ
3194348	3197568	-	3	citE, caiB, caiA
3208650	3214976	-	6	Hsero_2813, tldD, tldE, Hsero_2816, exbD, exbB
3215031	3218208	-	2	Hsero_2819, prhA
3219600	3221066	+	2	prhI, Hsero_2823
3222761	3223996	-	2	Hsero_2825, trxA1
3225255	3229993	+	6	nifS, nifU, Hsero_2831, hesB1, Hsero_2833, fdx
3231401	3235333	+	5	nifW, fixA, fixB, fixC, fixX
3235415	3238045	-	3	modC, modB, modA
3238752	3242931	-	6	fdxB, Hsero_2846, Hsero_2847, nifX, nifN, nifE
3248199	3249200	+	2	Hsero_2855, Hsero_2856
3249525	3250717	-	2	Hsero_2858, fdx
3250791	3257563	-	11	Hsero_2860, Hsero_2861, fixU, nifS2, Hsero_2864, nifZ1, nifZ, Hsero_2867, hesB, fdxN, nifB
3260788	3262595	+	2	Hsero_2873, modE1
3270661	3276050	-	4	HSERO_RS14440, tolC, hlyD, cyaB
3281384	3284884	+	4	Hsero_2884, Hsero_2885, Hsero_2886, Hsero_2887
3285961	3287066	-	2	lrgB, Hsero_2890
3287279	3288671	+	2	Hsero_2891, Hsero_2892
3289537	3290414	+	2	Hsero_2894, Hsero_2895
3290763	3294655	+	2	narK, nirB

Inicio	Fim	*Strand*	Nº de genes	Genes
3294802	3299217	+	3	nirD, Hsero_2899, nasA
3299680	3300994	-	2	Hsero_2903, Hsero_2904
3312723	3313495	-	2	Hsero_2916, Hsero_2917
3318236	3322443	-	4	hemY, hemX, hemD, hemC
3327028	3328299	-	2	Hsero_2927, Hsero_2928
3331440	3334687	+	4	Hsero_2932, Hsero_2933, Hsero_2934, Hsero_2935
3337452	3345535	-	2	pbpC, Hsero_2939
3348025	3351094	-	3	gpt, purA, hisZ
3351418	3355156	-	4	hflC, hflK, hflX, hfq
3355296	3357865	-	2	Hsero_2949, Hsero_2950
3357998	3360042	-	2	HSERO_RS14790, hisS
3360181	3362541	-	2	gcpE, Hsero_2954
3367106	3369121	-	2	rpoS, nlpD
3369349	3374325	-	4	pcm, surE, Hsero_2963, dinG
3374758	3376751	-	2	Hsero_2965, Hsero_2966
3384903	3388184	-	4	sdhB, sdhA, sdhD (HSERO_RS14910), sdhC
3390696	3392267	+	2	Hsero_2977, Hsero_2978
3396753	3399371	-	2	cheW, cheA
3406500	3408307	-	2	Hsero_2991, Hsero_2992
3408527	3410739	-	3	avrD, mmyX, acpD
3418368	3420239	-	2	Hsero_3002, rluD
3427637	3428455	+	2	Hsero_3010, Hsero_3011
3432528	3441483	-	8	Hsero_3015, Hsero_3016, cheB, cheA, cheW, cheR, Hsero_3021, Hsero_3022
3441733	3443016	-	2	Hsero_3023, serB
3446690	3447954	+	2	ispD, ispF
3451387	3454244	-	2	Hsero_3031, tsr
3460589	3462218	-	3	Hsero_3038, Hsero_3039, Hsero_3040
3496811	3501774	-	3	Hsero_3044, Hsero_3045, sunT
3506790	3508409	-	2	yedZ, Hsero_3051
3508607	3510702	+	2	czcR, czcS
3512279	3515904	+	4	Hsero_3055, Hsero_3056, cydA, cydB
3522282	3524692	-	2	mmsB, mmsA
3530080	3531068	+	2	rpoE, Hsero_3074
3531097	3533156	-	2	Hsero_3075, caiD
3539949	3542010	+	2	Hsero_3085, Hsero_3086
3542072	3545092	+	3	Hsero_3087, Hsero_3088, Hsero_3089
3545715	3546824	+	2	Hsero_3091, Hsero_3092
3549294	3551950	+	2	cheZ, cheA
3554001	3556218	+	2	cysI, Hsero_3100
3557156	3560600	+	4	cysD, cysN, cysG, Hsero_3105
3560679	3563337	-	3	Hsero_3106, Hsero_3107, Hsero_3108
3563489	3566046	+	2	pepA, Hsero_3110
3588260	3589830	-	2	Hsero_3130, Hsero_3131

Inicio	Fim	*Strand*	Nº de genes	Genes
3589985	3593042	+	2	Hsero_3132, Hsero_3133
3597344	3598501	-	2	rhtB, Hsero_3141
3599547	3601311	-	2	iscS1, Hsero_3145
3611548	3613685	-	2	Hsero_3153, selD
3623486	3626180	-	3	Hsero_3164, Hsero_3165, Hsero_3166
3627250	3630038	+	3	Hsero_3168, Hsero_3169, Hsero_3170
3631153	3632318	-	3	Hsero_3172, Hsero_3173, Hsero_3174
3634222	3640589	-	6	abgB, Hsero_3178, dppF1, dppD1, dppC1, dppB1
3649081	3653314	-	3	oprN, Hsero_3188, emrK
3660600	3662660	-	2	Hsero_3199, fixG
3662818	3665997	-	4	fixP, fixQ (HSERO_RS16045), fixO, fixN
3666142	3671155	-	4	fixS, fixI, Hsero_3206, hemN
3675217	3676610	-	2	argE, Hsero_3213
3683314	3684638	+	2	Hsero_3221, aat
3687206	3688263	+	2	chaC, Hsero_3227
3689181	3691477	-	2	pldB, Hsero_3232
3694526	3696450	-	3	Hsero_3235, rpoE, Hsero_3237
3696723	3700439	+	5	Hsero_3238, Hsero_3239, Hsero_3240, Hsero_3241, Hsero_3242
3703639	3705638	-	2	Hsero_3247, Hsero_3248
3734415	3737894	+	2	xdhB, xdhC
3740253	3744652	+	3	tolC, emrA, emrB
3751761	3753746	-	2	Hsero_3274, Hsero_3275
3765352	3766870	-	2	cvpA, Hsero_3288
3770564	3773450	-	3	trpB, trpF, truA
3788634	3791127	+	3	wrbA, Hsero_3307, Hsero_3308
3809956	3811534	-	2	Hsero_3321, Hsero_3322
3811666	3813437	-	2	Hsero_3323, Hsero_3324
3816032	3817807	+	2	argM, Hsero_3328
3817795	3821723	-	4	Hsero_3329, Hsero_3330, Hsero_3331, Hsero_3332
3829770	3831550	+	2	Hsero_3339, Hsero_3340
3831642	3835485	-	3	Hsero_3341, betA, Hsero_3343
3836758	3838843	+	2	Hsero_3345, Hsero_3346
3842108	3843570	+	2	Hsero_3351, Hsero_3352
3844508	3846242	+	3	Hsero_3354, Hsero_3355, Hsero_3356
3846349	3849488	-	2	dctD, dctB
3854111	3855615	+	2	thiL, pgpA
3857305	3859306	+	3	Hsero_3366, Hsero_3367, Hsero_3368
3868503	3869540	-	2	Hsero_3375, Hsero_3376
3870401	3876466	-	4	Hsero_3377, cheB, cheR, Hsero_3380
3877724	3879782	-	2	Hsero_3382, Hsero_3383
3880139	3883065	+	2	Hsero_3384, piuC
3884659	3886371	-	2	tauC, tauB
3888922	3890970	+	2	Hsero_3391, Hsero_3392

Inicio	Fim	*Strand*	Nº de genes	Genes
3892648	3898924	+	6	Hsero_3394, Hsero_3395, Hsero_3396, Hsero_3397, Hsero_3398, xylB
3899258	3903885	+	3	Hsero_3401, Hsero_3402, Hsero_3403
3906956	3911277	+	3	Hsero_3406, Hsero_3407, Hsero_3408
3911337	3913778	+	4	Hsero_3409, N_302, Hsero_3410, Hsero_3411
3915386	3915954	-	2	Hsero_3413, HSERO_RS17095
3921502	3925201	-	3	Hsero_3420, Hsero_3421, Hsero_3422
3930659	3934255	-	4	Hsero_3427, pqiB, pqiA, pqiA
3961882	3963074	-	2	chrA, chrA
3982874	3983784	-	2	umuC, umuD
3986292	3989763	-	2	Hsero_3472, spoVK
4007887	4009274	+	2	qor, Hsero_3500
4018726	4019887	-	2	Hsero_3509, Hsero_3510
4020266	4022961	-	3	Hsero_3511, Hsero_3512, glnQ
4023006	4024442	-	2	glnP, glnH
4024890	4029954	-	6	gltL, gltK, gltJ, pcaC, hisJ, Hsero_3521
4038610	4041990	+	2	Hsero_3528, Hsero_3529
4048085	4052132	+	3	fecI, fecR, Hsero_3535
4052413	4055449	+	2	Hsero_3536, fliD
4055493	4057486	+	2	flgE, Hsero_3539
4082297	4085181	+	4	hemF, nadD, Hsero_3561, Hsero_3562
4085237	4087311	+	2	maf, cafA
4088785	4093648	+	4	Hsero_3566, Hsero_3567, rfaJ, rfaJ
4101233	4103052	+	2	Hsero_3575, gltX
4118064	4119043	+	2	Hsero_3590, yjaB
4119120	4121029	-	2	Hsero_3592, Hsero_3593
4121114	4122797	-	3	Hsero_3594, Hsero_3595, Hsero_3596
4123522	4124670	-	2	Hsero_3597, Hsero_3598
4134435	4136485	-	2	nusA, Hsero_3606
4144934	4148142	-	2	Hsero_3613, alaS
4150158	4152124	+	2	Hsero_3618, Hsero_3619
4167290	4168886	+	3	Hsero_3635, Hsero_3636, Hsero_3637
4169190	4171863	+	2	narK1, narU
4172043	4179734	+	5	narG, narH, narJ, narI, moaA
4179858	4182501	+	2	narX, narL
4182562	4183397	+	2	Hsero_3647, Hsero_3648
4183489	4184885	+	2	Hsero_3649, pilW
4185029	4186450	+	3	Hsero_3651, Hsero_3652, Hsero_3653
4189304	4193835	-	5	Hsero_3657, Hsero_3658, Hsero_3659, Hsero_3660, Hsero_3661
4193951	4196818	-	3	Hsero_3662, nagH, nagG
4212665	4213560	-	2	recR, Hsero_3679
4217968	4220281	-	2	rfaE, ugdH
4220332	4221850	-	2	Hsero_3686, Hsero_3687

Inicio	Fim	*Strand*	Nº de genes	Genes
4222151	4224227	-	2	ihfB, rpsA
4226552	4228612	-	2	tyrA, Hsero_3693
4240233	4241955	+	3	Hsero_3701, Hsero_3702, Hsero_3703
4242018	4253218	+	11	Hsero_3704, Hsero_3705, Hsero_3706, Hsero_3707, Hsero_3708, hutH, Hsero_3710, Hsero_3711, Hsero_3712, Hsero_3713, fabF
4253685	4255663	+	2	fabG, fabF
4262979	4266357	+	3	Hsero_3724, Hsero_3725, Hsero_3726
4266593	4268925	+	2	Hsero_3727, Hsero_3728
4271794	4274486	+	2	tar, Hsero_3731
4277291	4279998	+	2	Hsero_3733, Hsero_3734
4280100	4285633	+	2	Hsero_3735, Hsero_3736
4286685	4287796	-	2	Hsero_3738, Hsero_3739
4287839	4289667	-	2	Hsero_3740, rsbU
4292565	4296394	-	2	Hsero_3745, Hsero_3746
4297936	4298897	-	2	Hsero_3748, Hsero_3749
4299670	4301540	-	2	Hsero_3751, kaiC
4304069	4307780	+	3	Hsero_3757, Hsero_3758, Hsero_3759
4308065	4308672	+	2	phd, doc
4311680	4314280	-	3	Hsero_3765, aroE, Hsero_3767
4314410	4316030	+	2	pcaR, Hsero_3769
4323470	4327009	-	2	kdpE, kdpD
4327190	4330032	-	2	kdpC, kdpB
4346958	4348938	-	2	upgC, ugpE
4351672	4353215	+	2	ubiA, tpm
4353264	4353955	-	2	Hsero_3798, Hsero_3799
4354239	4356078	-	3	Hsero_3800, nlpD, proC
4364147	4368814	-	3	Hsero_3811, Hsero_3812, Hsero_3813
4369087	4370668	-	2	fabG, fabG
4375883	4378574	+	3	fadD, Hsero_3822, Hsero_3823
4380559	4384398	-	3	glcF, glcE, glcD
4390150	4391372	-	2	ispA, xseB
4396865	4401994	-	3	Hsero_3839, pilJ, pilI
4402723	4404885	+	2	thiD, hemL
4405440	4407616	-	3	ccmA, Hsero_3847, argC
4407874	4408704	-	2	rpsI, rplM
4412649	4414097	-	2	gltK, gltJ
4433973	4436034	+	3	hisM, hisM, Hsero_3874
4438700	4440656	+	2	Hsero_3877, fabG
4454035	4457171	+	3	trpE, trpG, trpD
4461550	4463732	-	2	ubiB, Hsero_3901
4466329	4470786	-	2	hit, Hsero_3906
4478324	4480278	-	2	mutY, mutM
4480476	4482907	+	2	Hsero_3916, lolB

Inicio	Fim	*Strand*	Nº de genes	Genes
4507684	4508974	-	2	cybB, acpD
4517574	4524145	-	6	cheD, Hsero_3944, mhpA, HSERO_RS19705, Hsero_3946, Hsero_3947
4524212	4525762	-	2	Hsero_3948, Hsero_3949
4538863	4540762	+	2	Hsero_3958, Hsero_3959
4542168	4543904	-	2	ssb, Hsero_3962
4554244	4556431	+	3	Hsero_3968, Hsero_3969, Hsero_3970
4556501	4557851	+	2	Hsero_3971, Hsero_3972
4560263	4561200	+	2	Hsero_3975, ubiX
4572141	4573365	+	2	pcaH, pcaG
4578251	4580266	-	2	vanB, vanA
4583639	4586982	-	4	livF, livG, livM, livH
4588457	4589769	-	2	pcaJ, pcaI
4595210	4599567	+	4	Hsero_4007, paaJ, Hsero_4009, Hsero_4010
4601066	4604648	+	3	Hsero_4012, Hsero_4013, Hsero_4014
4604707	4607510	+	3	Hsero_4015, aidB, Hsero_4017
4609602	4613997	-	4	Hsero_4020, Hsero_4021, Hsero_4022, Hsero_4023
4619344	4624940	+	3	Hsero_4027, Hsero_4028, Hsero_4029
4628948	4632333	-	3	Hsero_4032, Hsero_4033, Hsero_4034
4632455	4634985	-	3	Hsero_4035, Hsero_4036, Hsero_4037
4636735	4638857	+	2	Hsero_4040, Hsero_4041
4643638	4646443	-	3	petC, petB, petA
4647995	4650634	+	3	degQ, Hsero_4053, glcC
4656255	4658650	-	4	hisE, hisI, hisF, hisA
4660097	4663341	-	3	hisC, Hsero_4068, hisD
4663386	4667282	-	5	hisG, murA, HSERO_RS20360, Hsero_4072, Hsero_4073
4667589	4668552	-	2	Hsero_4074, Hsero_4075
4668768	4669981	-	2	vacJ, Hsero_4077
4670044	4671551	-	2	Hsero_4078, Hsero_4079
4677072	4681934	+	2	Hsero_4085, Hsero_4086
4682828	4684469	+	2	Hsero_4089, Hsero_4090
4695226	4696931	+	2	hipA, Hsero_4100
4699965	4704062	+	3	Hsero_4103, Hsero_4104, Hsero_4105
4705044	4708758	+	2	exuT2, Hsero_4108
4711191	4715495	+	5	Hsero_4111, accB, accC, Hsero_4114, Hsero_4115
4718279	4721938	-	3	livF, livGM, livH
4728335	4729927	+	3	paaB, paaC, paaD
4729991	4735596	+	5	paaE, Hsero_4130, paaG, paaH, paaI
4735646	4739013	+	3	paaK, pacF1, Hsero_4136
4742787	4743980	-	2	kdtB, Hsero_4142
4744027	4747425	-	3	Hsero_4143, Hsero_4144, Hsero_4145
4747817	4750399	+	3	ftsY, ftsE, ftsX
4752953	4755574	-	3	Hsero_4151, ctaB, ctaA

Inicio	Fim	*Strand*	Nº de genes	Genes
4758370	4762651	-	6	Hsero_4158, coxG, N_364, coxA, coxB, Hsero_4162
4764130	4765549	+	2	comFC, cspR
4765611	4767300	-	2	Hsero_4166, Hsero_4167
4776849	4778890	+	2	Hsero_4177, Hsero_4178
4781509	4783196	+	2	Hsero_4180, Hsero_4181
4783761	4786390	+	2	Hsero_4182, pilQ
4786592	4789535	+	2	aroBK, dgt
4798122	4803210	+	5	wecC, galE, Hsero_4195, Hsero_4196, wecE
4803297	4807350	+	3	Hsero_4198, Hsero_4199, asnB
4807398	4813863	+	9	Hsero_4201, mesJ, hisH, hisF, Hsero_4205, lpxD, Hsero_4207, fabH, acpP
4816844	4828107	+	9	asnB, gumH, Hsero_4214, asnB, Hsero_4216, rfaG, rfbD, galT, wecB
4835402	4836058	+	2	Hsero_4228, Hsero_4229
4838683	4841545	-	3	Hsero_4232, hop, Hsero_4234
4841755	4842681	-	2	fadR, Hsero_4236
4845823	4848266	+	2	Hsero_4240, Hsero_4241
4848316	4854468	-	5	Hsero_4242, Hsero_4243, Hsero_4244, Hsero_4245, Hsero_4246
4854537	4861566	-	2	Hsero_4247, Hsero_4248
4869573	4871789	+	2	Hsero_4256, Hsero_4257
4878161	4879752	+	2	Hsero_4262, Hsero_4263
4879939	4883519	+	3	Hsero_4264, Hsero_4265, Hsero_4266
4884833	4885526	-	2	Hsero_4268, Hsero_4269
4886521	4887607	+	2	minD, minE
4891285	4892612	+	2	dsbB, hemK
4898851	4901534	-	3	Hsero_4281, cca, gst
4905309	4906507	-	2	metF, Hsero_4288
4909735	4912363	+	3	htrB, htrB, dapF
4921853	4923399	-	2	Hsero_4303, birA
4943658	4949324	+	4	vgr, Hsero_4322, Hsero_4323, Hsero_4324
4963968	4964926	-	3	Hsero_4337, rnpA, rpmH
4968102	4971572	+	3	Hsero_4341, ycgB, Hsero_4343
4974984	4976089	+	2	Hsero_4347, Hsero_4348
4981098	4982891	-	2	fmt, Hsero_4352
4983786	4986200	-	2	ftsN, argS
4995107	4997551	-	2	atpG, atpA
4997597	4998602	-	2	atpH, atpF
4998998	5000160	-	2	atpB, atpI
5005449	5008203	+	2	Hsero_4374, Hsero_4375
5008274	5009785	+	2	Hsero_4376, dadA
5009960	5013797	-	3	Hsero_4378, Hsero_4379, Hsero_4380
5014046	5016204	+	2	ompR, Hsero_4382
5017063	5019269	-	2	Hsero_4384, Hsero_4385
5021619	5022799	-	3	Hsero_4389, Hsero_4390, Hsero_4391

Inicio	Fim	*Strand*	Nº de genes	Genes
5028754	5030741	-	2	Hsero_4396, Hsero_4397
5032202	5033382	+	2	fabG, folB
5033485	5034212	+	2	Hsero_4401, Hsero_4402
5034299	5036114	-	2	Hsero_4403, tyrB
5040212	5042795	+	3	galE, rfbB, rfbC
5043316	5045303	+	2	Hsero_4413, wcaG
5045370	5047852	+	3	Hsero_4415, Hsero_4416, Hsero_4417
5048017	5049766	+	2	Hsero_4418, Hsero_4419
5051041	5053058	-	2	Hsero_4421, wcaA
5059892	5063429	-	2	Hsero_4429, Hsero_4430
5064419	5065477	-	2	Hsero_4431, Hsero_4432
5065641	5070593	-	3	Hsero_4433, czcA, czcB
5076754	5082215	-	5	fucK, Hsero_4443, Hsero_4444, Hsero_4445, Hsero_4446
5090398	5095261	-	5	serA, menG, Hsero_4455, Hsero_4456, Hsero_4457
5097193	5098728	+	2	Hsero_4460, menG
5110141	5115176	-	2	gatA, Hsero_4471
5115286	5118258	-	4	Hsero_4472, Hsero_4473, Hsero_4474, Hsero_4475
5122442	5125682	+	2	Hsero_4479, Hsero_4480
5125857	5127343	+	2	fecI, fecR
5130837	5131784	-	2	Hsero_4485, acpD
5136868	5137745	-	2	pinR, Hsero_4491
5137969	5140358	-	2	galM, xylR
5141747	5144486	+	2	xylG, xylH
5144750	5148287	+	3	fabG, Hsero_4498, Hsero_4499
5174019	5175585	-	2	Hsero_4513, fecI
5178208	5180527	-	2	Hsero_4518, Hsero_4519
5181618	5182037	+	2	Hsero_4521, Hsero_4522
5196666	5198797	+	2	ompR, tctE
5208341	5209379	+	2	hemO, Hsero_4545
5212154	5214232	-	2	Hsero_4548, Hsero_4549
5216160	5222207	+	6	Hsero_4551, Hsero_4552, Hsero_4553, nagK, nagA, glmS
5225775	5232792	-	3	Hsero_4559, Hsero_4560, Hsero_4561
5234638	5237693	+	2	betA, Hsero_4565
5242230	5244122	-	2	Hsero_4571, Hsero_4572
5244315	5247668	-	3	Hsero_4573, Hsero_4574, Hsero_4575
5247734	5251556	-	3	nadE, pncB, Hsero_4578
5264507	5266793	+	2	Hsero_4588, toeA
5273347	5275421	-	2	Hsero_4596, Hsero_4597
5276464	5277244	-	2	Hsero_4599, Hsero_4600
5281843	5284104	-	3	Hsero_4605, Hsero_4606, Hsero_4607
5287330	5293276	-	3	Hsero_4610, Hsero_4611, Hsero_4612
5293532	5295690	+	2	baeS, baeR

Inicio	Fim	*Strand*	Nº de genes	Genes
5297892	5303764	-	3	tolC, Hsero_4617, Hsero_4618
5305048	5306870	-	2	Hsero_4620, Hsero_4621
5306938	5308577	-	3	Hsero_4622, Hsero_4623, HSERO_RS23120
5309513	5311089	-	2	Hsero_4626, Hsero_4627
5316483	5319367	-	2	Hsero_4634, Hsero_4635
5320979	5322338	-	2	atoA, atoD
5332623	5335472	+	2	Hsero_4647, Hsero_4648
5337460	5340828	+	4	livH, livM, livG, livF
5340997	5343794	+	2	Hsero_4654, Hsero_4655
5352569	5356189	+	3	mreC, mreD, mrdA
5359714	5361381	+	2	Hsero_4669, Hsero_4670
5361504	5362752	+	2	gmhA, Hsero_4672
5378219	5379231	+	2	nahD, Hsero_4689
5379309	5383921	+	3	Hsero_4690, accC, Hsero_4692
5401096	5403236	-	2	Hsero_4708, Hsero_4709
5403542	5406700	+	2	Hsero_4710, Hsero_4711
5409498	5413824	+	4	urtB, urtC, urtD, urtE
5415080	5418670	+	5	ureA, ureB, ureC, ureE, ureF
5420023	5424036	-	4	glpD1, Hsero_4727, Hsero_4728, Hsero_4729
5425467	5427559	+	2	Hsero_4731, Hsero_4732
5434673	5438076	-	3	dctD1, Hsero_4740, Hsero_4741
5442747	5443939	-	2	Hsero_4745, folE
5446748	5449059	-	2	Hsero_4751, Hsero_4752
5450146	5452806	-	2	Hsero_4754, Hsero_4755
5468640	5471723	+	2	Hsero_4768, Hsero_4769
5471730	5473490	-	2	Hsero_4770, lipA
5473531	5474648	-	2	lipB, Hsero_4773
5476395	5477580	+	2	Hsero_4776, Hsero_4777
5490855	5492276	-	2	parA, rhtB
5496927	5498804	-	2	Hsero_4792, Hsero_4793
5500333	5501861	-	2	Hsero_4795, Hsero_4796

Anexo 07: *Script* em Perl para trimar os arquivos de qualidade.

```
#!/usr/bin/perl  -w
# Escrito por Vinicius Weiss - UFPR - 2012

$file = shift or die "Usage: meu.pl <fasta>\n";
$out=shift or die "arquivo de saida\n";
$seed=shift or die "\n especifique quanto deseja cortar";
open (AMOSTRA , "<$file");
open OUT, (">$out");
%reads;
while ($line=<AMOSTRA>){
  chomp $line;
    if ($line=~/#/) {

  #print ARQ "$line\n";
    }elsif ($line=~/\>/){
     $id=$line;
   }else{
     $reads{$id}=$line;
   }
}
print "ETAPA 1 --OK\n";
foreach $key (sort keys %reads) {
  $seq= $reads{$key};
  $tam=length($seq);
  ($seq)=(substr($seq,0,($tam-$seed+1)));
  print OUT "$key\n$seq\n";
}
```

Anexo 08: *Script* em Perl para trimar os arquivos csfasta.

```
#!/usr/bin/perl  -w
# Escrito por Vinicius Weiss - UFPR - 2012

$file = shift or die "Usage: meu.pl <fasta>\n";
$out=shift or die "arquivo de saida\n";
$seed=shift or die "\n especifique quantas bases deseja cortar";
open (AMOSTRA , "<$file");
open OUT, (">$out");
%reads;

while ($line=<AMOSTRA>){

  chomp $line;
    if ($line=~/#/) {

    #print ARQ "$line\n";
    }elsif ($line=~/\>/){
        $id=$line;
    }else{
        $reads{$id}=$line;
    }
}
print "ETAPA 1 --OK\n$tam\n";
foreach $key (sort keys %reads){
    $seq= $reads{$key};
    @vecSeq = split(' ',$seq);
    $tam=@vecSeq;
    (@newSeq)=(@vecSeq[0..($tam-$seed)]);
    $seq=join(' ',@newSeq);
    print OUT "$key\n$seq\n";

}
```