

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE MICROBIOLOGIA

Tese

Anne Cybelle Pinto

**ANÁLISE EM LARGA ESCALA DA EXPRESSÃO
DIFERENCIAL DE *Corynebacterium pseudotuberculosis* EM
RESPOSTA A ESTRESSES ABIÓTICOS**

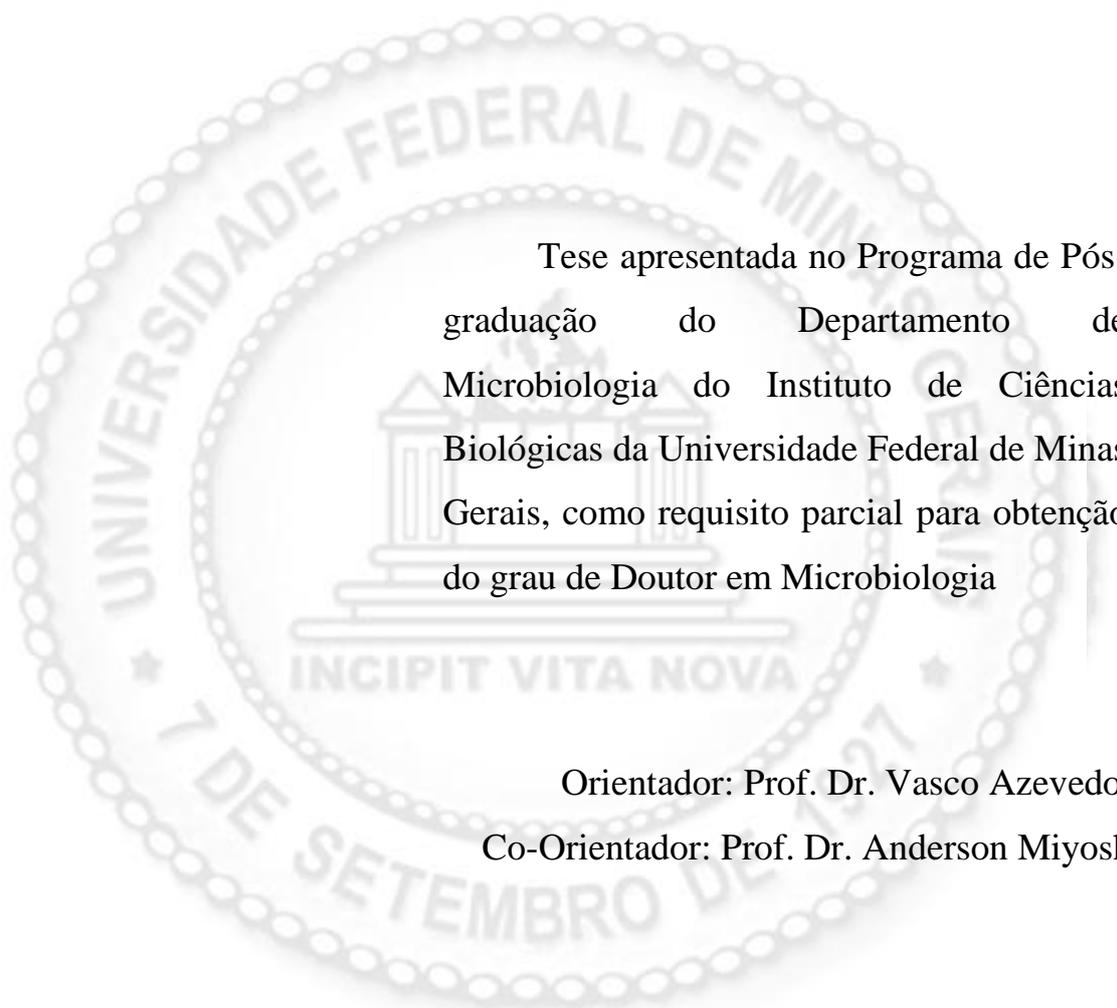


Belo Horizonte

2011

Anne Cybelle Pinto

**ANÁLISE EM LARGA ESCALA DA EXPRESSÃO
DIFERENCIAL DE *Corynebacterium pseudotuberculosis* EM
RESPOSTA A ESTRESSES ABIÓTICOS**



Tese apresentada no Programa de Pós-graduação do Departamento de Microbiologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do grau de Doutor em Microbiologia

Orientador: Prof. Dr. Vasco Azevedo
Co-Orientador: Prof. Dr. Anderson Miyoshi

Belo Horizonte

2011

Dedico este trabalho a minha mãe Dorinha, vovó Dodora (minha segunda mãe), irmãs Ariadna e Milene por todo incentivo, compreensão, por não me deixarem desistir de nada. Aos meus tesouros Ester e Rafaela. Dedico ainda ao meu noivo João Gomide pelo apoio constante, paciência, compreensão, amor. Amo vocês!!!!

"O conhecimento a gente adquire com os mestres e os livros. A sabedoria com a vida e com os humildes."

Autor desconhecido

"Um pouco de ciência nos afasta de Deus. Muito, nos aproxima"

Louis Pasteur

AGRADECIMENTOS

Sempre, em primeiro lugar a Deus por tudo que me acontece

As Universidades: Federal de Minas Gerais e Federal do Pará por permitirem meu crescimento profissional;

A todos do Instituto de Ciências Biológicas (faxineiros, porteiros e outros funcionários), obrigada por manter (ou tentar manter) o ambiente mais agradável;

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa, por acreditarem neste projeto;

Ao prof Dr. Vasco Azevedo por me dar a oportunidade de ingressar ao doutorado. Obrigada pela orientação e ensinamentos. Obrigada por ter dado a oportunidade de conhecer a UFPA e o Laboratório de Polimorfismo de DNA. Aprendi muito o tempo que estive lá. Valeu muito a pena;

Ao prof Anderson Miyoshi por todos os ensinamentos e disponibilidade em todos os momentos que precisei;

A TODOS do LGCM, obrigada pela convivência e amizade, pela paciência, ensinamentos, pela compreensão quando “tomei o NB2” só pra mim...Foi necessário...;

Ao prof Dr. Artur Silva por não medir esforços em ajudar e pela acolhida no laboratório (LPDNA) em Belém do Pará. Obrigada pela paciência, por me “co-orientar”, por não deixar faltar nada para que o resultado fosse realmente excelente. Por estar sempre pronto em me receber quando precisei. Muito obrigada!!!

A Silvanira Barbosa, técnica do LPDNA, muito obrigada por tudo!!!!!! VOCÊ É MUITO ESPECIAL. EXCELENTE PROFISSIONAL!!! Tirei a sorte grande quando me colocaram para fazer o trabalho contigo;

A TODOS do LPDNA por terem me recebido de maneira extraordinária. Muito obrigada pela convivência, ensinamentos, disponibilidade (sempre, sem noção). Adorei a convivência com TODOS sinceramente;

Em especial, agradeço ao Pablo Gomes de Sá, um dos bioinformatas “de plantão”, que foi direcionado a me auxiliar e contribuiu além do esperado;

Ao Rommel Ramos, este agradeço de maneira mais que especial. Obrigada por abrir as portas da sua casa em Belém para me acolher, literalmente. Conforto impar, atenção, carinho, dicas, amizade. Obrigada por aliviar a saudade da minha família com uma convivência tão MARAVILHOSA!!!!!! Além de tudo, um excelente bioinformata que também não dei sossego

e me salvou de muitas. Foi ótimo passarmos a madrugada desenvolvendo trabalho, corrigindo paper, conversando e etc e etc e etc...VALEU DEMAIS;

A Ana Ramos, mãe do Rommel, por me receber sem conhecer (correu risco rrsrrsrrs), dando todo conforto e preocupada em não faltar nada para mim. Obrigada pelos almoços deliciosos, pelas conversas, pela companhia. Aprendi muito;

Aos grandes amigos Silvia, Eliane e Uiraci por todas as sugestões, conselhos, incentivo.....muito obrigada;

Obrigada a todos os meus familiares e amigos (obrigada M.A.E)

RESUMO	7
ABSTRACT	8
1.1. Características de <i>Corynebacterium pseudotuberculosis</i>	9
1.1.1. Aspectos microbiológicos e bioquímicos	9
1.1.2. Susceptibilidade antimicrobiana	9
1.1.3. Taxonomia	10
1.1.4. Determinantes de virulência	11
1.1.4.1. Fosfolipase D (PLD)	11
1.1.4.2. Lipídeos associados à parede celular	12
1.2. A Doença	12
1.2.1. Manifestação Clínica e Patogenia	12
1.2.2. Epidemiologia	13
1.2.3. Transmissão	14
1.2.4. Tratamento e Profilaxia	14
1.2.5. Diagnóstico	16
1.2.6. Vacinas	17
1.3. Estresses Abióticos	18
1.3.1. Estresse osmótico	19
1.3.2. Estresse por acidez	22
1.3.3. Estresse térmico	25
1.4. Projeto Genoma de <i>Corynebacterium pseudotuberculosis</i>	28
1.5. Plataforma SOLiD™	30
1.5.1. Uso do SOLiD™ para a análise do transcriptoma	31
2. JUSTIFICATIVA	39
3. OBJETIVO	40
3.1. Objetivos específicos	40
4. MATERIAL E MÉTODOS –ETAPA TRANSCRIPTOMA	41

4.1 Linhagem utilizada	41
4.2 Curva de crescimento da linhagem 1002	41
4.3 Avaliação da susceptibilidade de <i>C. pseudotuberculosis</i> aos estresses	41
4.3.1 Experimentos de resistência ao estresse térmico	41
4.3.2 Experimentos de resistência ao estresse osmótico	42
4.3.3 Experimentos de resistência ao estresse por acidez	42
4.4. Obtenção do material para extração do RNA	42
4.5. Extração do RNA total	43
4.6. Remoção do rRNA e obtenção do cDNA	43
4.7. Preparo da biblioteca de cDNA	44
4.8 Preparo das microesferas para o sequenciamento	46
4.9. Análise dos dados	47
4.10. Análise da expressão diferencial e seleção dos genes diferencialmente expressos	48
5. RESULTADOS	50
5.1. Curva de Crescimento	50
5.2. Avaliação da susceptibilidade de <i>C. pseudotuberculosis</i> aos estresses	50
5.3. Extração do RNA	51
5.4. Processamento dos dados do SOLiD™ e Análise em Softwares	52
5.5. Quantidade de transcritos representados em cada condição testada	56
5.6. Genes envolvidos na virulência	58
5.6.1. Gene da <i>pld</i> e genes <i>fag</i>	58
5.6.2. Gene <i>cp40</i>	62
5.6.3. Componentes da parede celular e modificação lipídica	63
5.7. Core estimulon entre as três condições de estresse, com <i>fold-change</i> ≥ 2 ou $\leq 0,5$	65
5.8. Estimulon no estresse osmótico, com <i>fold-change</i> ≥ 2 ou $\leq 0,5$	67

5.8.1. Processo Biológico do estimulon dos genes induzido no meio osmótico	73
5.8.2. Processo Biológico do estimulon dos genes reprimidos no meio osmótico	73
5.9. Estimulon do estresse por acidez, com <i>fold-change</i> ≥ 2 ou $\leq 0,5$	74
5.9.1. Processo Biológico do estimulon dos genes induzidos na acidez	78
5.9.2. Processo Biológico do estimulon dos genes reprimidos na acidez	79
5.10. Estimulon do estresse térmico, com <i>fold-change</i> ≥ 2 ou $\leq 0,5$	80
5.10.1. Processo Biológico do estimulon dos genes induzidos no choque térmico	81
5.10.2. Processo Biológico do estimulon dos genes reprimidos no choque térmico	82
5.11. Expressão dos Genes codificantes de fatores sigmas	83
5.12. Ilhas de Patogenicidade	84
5.13. Produtos gênicos confirmados, <i>in silico</i> ou experimental, por outros trabalhos do grupo LGCM, envolvendo <i>C.pseudotuberculosis</i> 1002	88
6. DISCUSSÃO	89
6.1. Análise da expressão dos genes mais caracterizados em <i>C. pseudotuberculosis</i>	90
6.1.1. Fosfolipase D	90
6.1.2. Genes <i>fag</i>	92
6.1.3. Serina protease	94
6.1.4. Componentes da parede celular e modificação lipídica	96
6.2. Core Estimulon entre os três estresses, com <i>fold-change</i> ≥ 2 ou $\leq 0,5$	102
6.3 Estimulon no estresse osmótico	116
6.3.1. Genes induzidos	116
6.3.2. Genes reprimidos	118
6.3.3 Processos Biológicos do Estimulon Osmótico	120
6.3.3.1. Processo Induzido	120
6.3.3.2. Processo Reprimido	121
6.4. Estimulon da acidez	122
6.4.1 Genes induzidos	122
6.4.2. Genes reprimidos	125
6.4.3. Processo Biológico do Estimulon no meio ácido	128
6.4.3.1. Processo Induzido	128
6.4.3.2. Processo Reprimido	129

6.5. Estimulon do choque térmico	130
6.5.1. Genes induzidos	130
6.5.2. Genes reprimidos	134
6.5.3. Processo Biológico do Estimulon no meio térmico	136
6.5.3.1. Processo Induzido	136
6.5.3.2. Processo Reprimido	137
6.6. Genes codificantes dos fatores sigmas	137
6.7. Genes das Ilhas de Patogenicidade preditas no genoma de <i>C.pseudotuberculosis</i> 1002	146
6.8 Produtos gênicos confirmados, in silico ou experimental, por outros trabalhos do grupo LGCM, envolvendo <i>C.pseudotuberculosis</i> 1002	153
6.9. Ilustração de um possível modelo de regulação entre alguns genes diferencialmente expressos	154
7. CONCLUSÕES	156
8. PERSPECTIVA	157
9. REFERÊNCIAS BIBLIOGRÁFICAS	158
10. ANEXOS	188

Lista de figura

FIGURA 1 - Quatro cores utilizadas no sistema SOLiD™ para detectar as duas bases, em 16 possíveis combinações.	30
FIGURA 1- <i>Pipeline</i> resumindo os passos para preparo das bibliotecas no SOLiD™	45
Figura 2: Descrição dos passos para preparo da PCR em emulsão	46
FIGURA 4- <i>Pipeline</i> do BioScope.	48
FIGURA 5- Curva de crescimento de <i>C.pseudotuberculosis</i> linhagem 1002 em condição controle, medida em densidade óptica de 600 nm (DO _{600nm}).	50
FIGURA 6- Número de células viáveis em cada condição.	51
FIGURA 7 - RNA total, tratado com DNase.	52
FIGURA 8 - RNA depletado dos ribossomais.	52
FIGURA 9 - Deposição das <i>beads</i> no SOLiD™ em cada condição testada.	53
FIGURA 10 - Gráfico representando a distribuição da média de qualidade em cada base do transcrito (50 nt) em cada condição.	54
FIGURA 11 - Número total de transcritos ativos que foram confirmados pelo RNAseq através do programa Bioscope.	57
FIGURA 12 – Quantidade de genes considerados diferencialmente expressos em relação ao controle.	58
FIGURA 13 - Cobertura da região transcricionalmente ativa do gene <i>pld</i> e <i>fagD</i> .	59
FIGURA 14 - Cobertura da região transcricionalmente ativa do gene <i>fagA</i> e <i>fagD</i> .	60
FIGURA 15 - Regulon do gene <i>dtxR</i> .	61
FIGURA 16 - Genes envolvidos na regulação da síntese de lipídeos.	64
FIGURA 17 - Frequências de genes induzidos pertencentes ao estímulo osmótico.	73
FIGURA 18 - Frequências de genes reprimidos pertencentes ao estímulo osmótico.	74
FIGURA 19 - Frequências de genes induzidos pertencentes ao estímulo ácido.	79
FIGURA 20 - Frequências de genes reprimidos pertencentes ao estímulo ácido.	79
FIGURA 21 - Cobertura dos genes (operon) <i>dnaK-grpE-dnaJ-hspR</i> , envolvidos na resposta ao choque térmico. Genes expressos no meio sob estresse térmico à esquerda e no controle, à direita. Observa-se a diferença na expressão entre as duas condições. Obtida pelo Artemis através do arquivo de saída .bam do Bioscope.	81
FIGURA 22 - Frequências de genes induzidos pertencentes ao estímulo térmico.	82
FIGURA 23 - Frequências de genes reprimidos pertencentes ao estímulo térmico.	82
FIGURA 24 - Mapa de cores relacionado ao nível de expressão dos genes codificantes dos fatores sigmas.	84
FIGURA 25 - Genes diferencialmente expressos presentes nas ilhas de patogenicidade de <i>C.pseudotuberculosis</i> 1002.	87
FIGURA 26 - Biossíntese da lipoproteína bacteriana.	100
FIGURA 27 – Representação de uma possível regulação entre os genes diferencialmente expressos	155

Lista de Tabela/Quadro

QUADRO 1: Linhagens que foram sequenciadas (exceto C231, 106, 399 e 4202), montadas e anotadas até o momento pelo grupo LGCM em parceria com RGMG	23
TABELA 1: Quantidade de leituras totais obtidas no sequenciamento	55
TABELA 2 - Quantidade de leituras únicas mapeadas no genoma de <i>C.pseudotuberculosis</i> 1002 e cobertura dos transcritos no genoma.	55
TABELA 3 - Expressão diferencial, em cada condição, do gene <i>pld</i> , genes do operon <i>fagABC</i> e gene <i>fagD</i> .	58
TABELA 4 - Expressão diferencial do gene <i>dtxR</i> .	60
TABELA 5 - Expressão de genes supostamente codificantes de serina protease.	62
TABELA 6 - Genes codificantes de proteínas componentes da parede.	63
TABELA 7 - Genes envolvidos na síntese de lipídeos.	64
TABELA 8 - Genes induzidos, comuns, nas três condições de estresse, com <i>fold-change</i> maior ou igual a dois.	66
TABELA 9 - Genes reprimidos comuns entre os três estresses	67
TABELA 10 - Genes induzidos na condição de estresse osmótico, com <i>fold-change</i> ≥ 2	68
TABELA 11 - Genes reprimidos na condição de estresse osmótico, com <i>fold-change</i> $\leq 0,5$.	71
TABELA 12 - Genes induzidos na condição de estresse ácido, com <i>fold-change</i> ≥ 2 .	74
TABELA 13 - Genes reprimidos na condição de acidez, com <i>fold-change</i> $\leq 0,5$	76
TABELA 14 - Genes induzidos na condição de estresse térmico, com <i>fold-change</i> ≥ 2 .	80
TABELA 15 - Genes reprimidos no estresse térmico, com <i>fold-change</i> $\leq 0,5$.	81
TABELA 16 - Expressão dos genes codificantes dos fatores sigmas.	83
TABELA 17 - Genes presentes nas ilhas de patogenicidade que foram considerados diferencialmente expressos	84

RESUMO

Uma nova tecnologia de sequenciamento de cDNA tem permitido uma análise detalhada e aprofundada dos transcritos em procariotos. *Corynebacterium pseudotuberculosis*, uma bactéria causadora de doenças, como a linfadenite caseosa, que trazem grandes prejuízos econômicos para o agronegócio mundial, foi o alvo de nossos estudos. Neste contexto, caracterizamos o perfil transcricional diferencial deste microrganismo em algumas condições que simulam o ambiente encontrado pela bactéria no hospedeiro no processo de infecção como, estresse osmótico (2M), térmico (50°C) e acidez (pH=5). No presente estudo escolhemos a plataforma SOLiD™ para tentar compreender o desenvolvimento da doença e a modulação da resposta a partir dos genes envolvidos no processo da infecção. Conseguimos detectar todas as regiões transcionalmente ativas preditas no genoma (2057 genes), sendo que na condição controle (condição fisiológica normal), foi possível observar que 97,58% dos genes são transcritos, apresentando 2,42% de região não transcrita com valor de RPKM (Reads Per Kilobase of coding sequence per Million mapped) igual a zero. Nas condições de estresse osmótico, estresse térmico e acidez, 98,05%, 97,34%, 97,81% dos genes são transcritos, respectivamente. Dos 2064 transcritos do estresse osmótico, 48,01% foram considerados diferencialmente expressos em relação ao controle, 45,68% genes considerados diferencialmente expressos no térmico e no estresse ácido, 48,08%. Grande parte dos genes em todas as condições, induzidos ou reprimidos foi composta por proteínas de funções desconhecidas (hipotéticas) confirmando quão pouco este microrganismo é caracterizado. As análises revelaram uma forte correlação entre resposta ao estresse ácido e expressão de genes relacionados à virulência. O processo de regulação da transcrição e adesão celular estão entre os principais resultados do Blast2GO. No estresse osmótico processo de reparo e transporte transmembrana foram destaques, além do processo de adesão. No estresse térmico, menor número de genes diferencialmente expressos fizeram parte da análise e entre os principais processos estão o de oxidação e catabólico de ATP, gerando energia para célula sobreviver no ambiente hostil. Dentro do *core estimulon* (genes compartilhados entre os estresses), genes de virulência se destacaram, demonstrando que já no início da fase exponencial a bactéria regula a transcrição para uma resposta específica e adaptativa. Entre os reprimidos, os genes envolvidos no processo metabólico e biossintéticos foram os mais representados, demonstrando redução no crescimento, que é uma importante estratégia de ajuste da fisiologia celular para uma nova condição. O presente estudo permitiu criar um catálogo gênico que é utilizado como recurso para sobrevivência dentro de ambientes prejudiciais e conseguir escapar da resposta imune do hospedeiro. Estes resultados nos permitem sugerir possíveis candidatos à obtenção de vacinas eficientes com potencial redução das perdas no agronegócio, em estudos futuros.

Palavras-chaves: *Corynebacterium pseudotuberculosis*, linfadenite caseosa, expressão gênica, SOLiD™, estresses

ABSTRACT

A new cDNA sequencing technology has allowed deep analysis of prokaryotic transcripts. *Corynebacterium pseudotuberculosis*, our study target, is a bacterium responsible for several diseases, such as caseous lymphadenitis that causes important economic losses in global agribusinesses. In this context, we characterized its differential transcriptional profile in some conditions that mimic the environment found during an infectious process, namely osmotic (2M), termic (50°C) and acid (pH=5) stresses. For this work, SOLiD™ platform was chosen to try to understand the development of the disease and modulation of the response from the genes involved in the infectious process. All transcriptionally active regions that were predicted in the genome (2057 genes) were detected, observing 97,58% of transcript genes under the control condition, and 2,42% of non-transcribed regions with a RPKM value (Reads Per Kilobase of coding sequence per Million mapped) equal to zero. During conditions of osmotic, termic and acid stresses, 98,05%, 97,34% and 97,81% of genes, respectively, were found to be transcribed. During osmotic stress, 48,01% of the transcripts were considered differentially expressed with relation to the control, whereas during termic and acid stress, 45,68% and 48,08% respectively were found. It was observed that most genes for all stresses were formed by proteins of unknown function (hypothetical), highlighting the need for more information about this microorganism. Our analysis revealed a strong relationship between acid stress response and gene expression of genes previously associated to virulence. The process of transcription regulation and cell adhesion are among the main results of Blast2GO. During osmotic stress, repair process and transmembrane transport, as well as adhesion process, were highlighted. During heat shock, fewer differentially expressed genes were part of the analysis and oxidoreduction processes and ATP catabolic process were among the main energy generators for cell survival in the hostile environment. Within the core estimulon (genes shared between the stresses), virulence genes stood out, demonstrating that already at the beginning of the exponential phase the bacteria regulates the transcription for a specific and adaptive response. Among the repressed genes, those involved in metabolic and biosynthesis were the most represented, demonstrating reduced growth, an important strategy for adjusting cell physiology to a new condition. The present study allowed the creation of a gene catalog that can be used as a resource for survival in adverse environments and escape of the host immune response. These results allow us to suggest, in future studies, possible candidates for efficient vaccines regarding potential loss reduction in agribusinesses.

Key words: *Corynebacterium pseudotuberculosis*, caseous lymphadenitis, gene expression, SOLiD™, stress

1. INTRODUÇÃO

1.1. Características de *Corynebacterium pseudotuberculosis*

1.1.1. Aspectos microbiológicos e bioquímicos

Corynebacterium pseudotuberculosis é uma bactéria pertencente à classe Actinobacteria, à qual pertencem também os gêneros *Corynebacterium*, *Mycobacterium*, *Nocardia* e *Rhodococcus*, constituindo o grupo CMNR. Esse grupo possui características em comum como: (i) parede celular composta principalmente de peptidoglicanos, arabinogalactano e ácidos micólicos, e (ii) o alto conteúdo G + C do genoma (47-74%) (Coelho, 2007).

Esse microrganismo é Gram-positivo, intracelular facultativo, exibindo formas pleomórficas, que variam desde cocóides a bastões filamentosos e medem 0,5 µm a 0,6 µm por 1 µm a 3 µm de tamanho. Este microrganismo não possui cápsula, não esporula, é imóvel e possui fímbrias (Dorella et al., 2006a).

O crescimento de *C. pseudotuberculosis* na superfície do ágar é inicialmente espaçado na superfície e, posteriormente, se organiza em colônias opacas, de crescimento concêntrico e de coloração creme alaranjado; já em meio líquido, desenvolve-se como depósitos granulares. As condições ótimas de crescimento são a 37°C em pH entre 7,0 e 7,2, e a anaerobiose é facultativa (Dorella et al., 2006a)

O peptidoglicano da parede celular apresenta, além de outros constituintes, o ácido meso-diaminopimélico que está relacionado à resistência da parede contra a maioria das peptidases. Arabinose e galactose são os principais açúcares da parede (Dorella et al., 2006a). As reações bioquímicas de isolados de *C. pseudotuberculosis* variam muito, principalmente em sua habilidade de fermentação. Todas as linhagens produzem ácido, mas não gás, a partir de fontes de carbono incluindo glicose, frutose, maltose, manose, e sacarose (Holt et al., 1994). As reações bioquímicas são catalase, urease e fosfolipase D positiva, oxidase negativa e beta hemolítica (Muckle e Gyles, 1982)

1.1.2. Susceptibilidade antimicrobiana

Com relação à susceptibilidade de *C. pseudotuberculosis* aos agentes antimicrobianos, ocorre uma variação dependendo da fonte de onde o isolado foi obtido. Uma maior resistência aos agentes antimicrobianos foi encontrada em análise de corinebactérias isoladas de mastite em ovelhas (Fernández et al., 2001). Muckle e Gyles (1982), em seus estudos com 26

linhagens isoladas de lesão de linfadenite caseosa (LC), relataram que todas foram susceptíveis aos antibióticos ampicilina, cloranfenicol, lincomicina, gentamicina, tetraciclina, penicilina G, sulfametoxazol/trimetopim. Três linhagens apenas foram susceptíveis a neomicina e todas foram resistentes a estreptomicina. A concentração inibitória mínima (CIM) para todos os isolados foi similar para vários agentes antimicrobianos. Porém, a resistência difere dependendo das condições infecciosas. Altos valores de CIM para diferentes agentes antimicrobianos foram observados em uma análise de corinebactérias isoladas de mastite em ovelhas (Fernández et al., 2001). Olson et al. (2002) relatam que *C. pseudotuberculosis* é altamente resistente a qualquer droga quando há a formação de biofilme, deixando o tratamento com antibiótico impraticável.

1.1.3. Taxonomia

A classificação taxonômica de *C. pseudotuberculosis* foi baseada primeiramente pelas características bioquímicas e morfológicas (Muckle e Gyles, 1982). Quanto à capacidade de redução de nitrato a nitrito, dois biovars foram identificados: tipo I, isolada, preferencialmente, de ovino ou caprino, nitrato negativo e tipo II, equino e bovino, nitrato positivo (Batey, 1986). Porém, outros autores têm relatado que ambos os tipos podem ser isolados de bovinos (Barakat et al., 1984). Ainda não está claro se a disseminação e doença visceral estão relacionadas à diferença na linhagem, aos fatores do hospedeiro, ou ambos (Shpigel et al., 1993). Para diferenciação entre gêneros de *Corynebacterium* e mesmo para distinguir, na espécie *C. pseudotuberculosis*, os dois biotipos, tipo I e tipo II, a técnica de análise de restrição de rDNA amplificado (ARDRA), que estuda o gene da subunidade 16S do rRNA bacteriano foi utilizada e mostrou-se muito eficiente (Vanechoutte et al., 1995). De acordo com estes estudos, *C. pseudotuberculosis* está filogeneticamente mais próximo de *C. ulcerans* do que de *C. diphtheriae*. A proximidade destes organismos foi sugerida pelo fato de serem os únicos entre corinebactérias a produzirem fosfolipase D (Buck et al., 1985).

Análises de sequenciamento do gene da subunidade β da RNA polimerase (*rpoB*) mostraram-se mais acuradas para a identificação das espécies de corinebactérias do que análises baseadas no rDNA 16S (Khamis et al., 2004). Dorella et al (2006a) demonstraram, através da árvore filogenética construída baseada na sequência de gene *rpoB*, a clara relação entre *C. pseudotuberculosis* e *C. ulcerans* mostrando que o gene *rpoB* é uma poderosa ferramenta de identificação. Porém, muitos autores propõem que as análises de subunidade β da RNA polimerase seja utilizada de forma complementar as análises do gene rRNA 16S para

estudos filogenéticos das espécies de *Corynebacterium* e *Mycobacterium* (Khamis et al., 2004; Kim et al., 1999).

1.1.4. Determinantes de virulência

Apesar de o processo patogênico da doença ser relativamente bem entendido, pouco se sabe sobre os determinantes moleculares de virulência da *C. pseudotuberculosis*, bem como sobre o controle da sua expressão gênica (Dorella et al., 2006a; Mckean et al., 2007). As bactérias patogênicas abrigam genes que permitem sua entrada inicial e a sobrevivência dentro do hospedeiro, bem como os genes que resultam na doença e são definidos como determinantes de virulência (Schumann, 2007). Billington et al. (2002) demonstraram que três genes presentes em um operon *fagABC* e gene *fagD* apresentam papel na virulência desta bactéria e foram designados como genes de aquisição de ferro. Outro determinante seria uma proteína de 40kDa, identificada como proteína imunogênica que revelou atividade proteolítica, como serina protease (Wilson, Brandon e Walker, 1995). Porém, apenas dois determinantes de virulência da *C. pseudotuberculosis* são melhores caracterizados: fosfolipase D e lipídeos da parede celular.

1.1.4.1. Fosfolipase D (PLD)

A fosfolipase D é caracterizada como uma exotoxina e muitas de suas propriedades biológicas ainda não estão esclarecidas. Algumas de suas características são conhecidas, tais como: capacidade de hidrolisar lisofosfatidilcolina e esfingomiéline; necessidade de íons cátiões Ca^{+} e Mg^{+} para sua atividade; capacidade de lisar eritrócitos de carneiro em sinergismos com a colesterol oxidase e a fosfolipase C produzidas por *Rhodococcus equi* (Songer, 1997). O gene da PLD de *C. pseudotuberculosis* já foi clonado e expresso em *Escherichia coli*. Codifica uma proteína de 31,4 KDa e está presente no genoma de *C. pseudotuberculosis* em cópia única (Hodgson et al., 1999; Songer et al., 1990).

Estudos utilizando a exotoxina bruta ou parcialmente purificada mostraram que a atividade da PLD sobre a esfingomiéline na membrana das células do endotélio vascular parece ser importante na disseminação da bactéria do local da infecção para os linfonodos, pois o uso de antitoxina limitou esta disseminação no hospedeiro infectado, no entanto não evitou o desenvolvimento de abscessos (Alves e Olander, 1999; Williamson, 2001). É sabido que a PLD atua na permeabilidade dos vasos linfáticos e sanguíneos e como uma hemolisina que tem ação necrosante (Hirsh e Zee, 2003). A disseminação da bactéria por ação desta exotoxina

ocorre pela ação dermonecrótica que a mesma exerce, danificando as células endoteliais e promovendo o extravasamento do plasma dos pequenos vasos sanguíneos da derme e, então, o acesso aos vasos linfáticos (Songer, 1997). A PLD é também considerada leucotóxica o que contribui para as lesões e a destruição dos macrófagos de caprinos durante infecção com *C. pseudotuberculosis* (Tashjian e Campbell, 1983).

1.1.4.2. Lipídeos associados à parede celular

Outro determinante de virulência de *C. pseudotuberculosis* é a presença de lipídeos associados à parede celular. Alto conteúdo de lipídeos confere ao microrganismo resistência à digestão por enzimas celulares e, conseqüentemente, a capacidade de persistir como parasita intracelular facultativo (Braithwaite et al., 1993). De acordo com Jolly (1966), esta camada lipídica dificulta a fagocitose da bactéria, aumentando sua virulência e linhagens mais virulentas possuem mais lipídeos que linhagens atenuadas. A toxicidade do material lipídico extraído foi demonstrada pela indução de necrose hemorrágica após injeção intradérmica em cobaias e de acordo com Muckle e Gyles (1982), há uma relação direta entre a porcentagem de lipídeos de superfície e a indução de abscessos crônicos. O lipídeo é também citotóxico e induz a necrose que determina o potencial de patogenicidade do microrganismo. Os efeitos são localizados e resultam na formação de abscessos (Collett et al., 1994), porém a extração desse lipídeo não afeta a viabilidade do microrganismo (Muckle e Gyles, 1982).

1.2. A Doença

1.2.1. Manifestação Clínica e Patogenia

C. pseudotuberculosis foi inicialmente identificada como microrganismo causador da patologia LC em ovelhas e cabras, porém também já foi isolada de outras espécies, como camelos, suínos, búfalos, bovinos, cavalos causando linfangite ulcerativa ou acne contagiosa, e até em humanos causando linfadenite subaguda a crônica (Williamson, 2001; Yeruham et al., 2004). São doenças infecto-contagiosas crônicas e caracterizam-se por um quadro de abscessos dos linfonodos que apresentam material purulento de cor branco-esverdeada (Ayers, 1977).

Ao invadir o hospedeiro, principalmente por feridas na pele, *C. pseudotuberculosis* é capturado por células fagocíticas. Porém o processo de fagocitose acaba sendo ineficiente, pois a bactéria resiste ao compartimento fagolisossoma, continua a sua multiplicação, causa lise das células fagocíticas e conseqüentemente pode infectar outros macrófagos (Jones et al., 2000).

Além disso, existe uma habilidade de formar biofilme o que evita a fagocitose por macrófagos e neutrófilos (Costerton, Stewart, Greenberg, 1999).

Essa incapacidade de eliminação da bactéria pode também ser devido à presença da camada lipídica e a possível ausência de óxido nítrico, que é um potente efector na eliminação de patógenos intracelulares, como *C.pseudotuberculosis* (Bogdan et al., 1997).

Com a morte das células do sistema imunológico, o microrganismo acomete o linfonodo drenante local ocasionando a formação de lesões necróticas (Jones et al., 2000). A partir do início de um novo ciclo de replicação, as bactérias conseguem colonizar os linfonodos e sobreviver, podendo atingir diversos órgãos e células.

Na tentativa de conter e eliminar o microrganismo ocorre a formação de granulomas. Microscopicamente a lesão se inicia pelo aparecimento de células epitelióides que, em seguida, são substituídas por necrose caseosa (Jones et al., 2000). A massa caseosa vai sendo circundada por células epiteliais e linfócitos e, em seguida por uma camada de tecido conjuntivo fibroso. À medida que a lesão progride, estas camadas celulares vão sofrendo necrose. Enquanto a camada fibrosa ainda é visível, o processo se repete sucessivamente. O processo crônico resulta no aspecto macroscópico desta patologia (Jones et al., 2000).

1.2.2. Epidemiologia

A doença tem sido identificada em países onde existem grandes criações de ovinos e caprinos como Austrália, Argentina, Nova Zelândia, África do Sul e Estados Unidos e em países da Comunidade Européia (França, Itália, Grã-Bretanha, União Soviética), Chile, Uruguai, Canadá, e Sudão. No Brasil, já foi isolada e caracterizada em vários Estados, tratando-se, portanto, de um microrganismo cosmopolita (Oliveira, 2007).

O Brasil é detentor de 3% do rebanho mundial de caprinos e ovinos. (EMBRAPA Semi-Árido, 2008). Na região nordeste, a incidência de LC foi verificada por (Unanian et al., 1985) onde examinaram 656 caprinos periodicamente durante dois anos e 41,6% dos animais apresentaram abscessos superficiais palpáveis. Os prejuízos provocados pela doença são muito grandes, uma vez que muitos dos pequenos criadores têm a caprinovinocultura como uma das suas principais atividades econômicas e os abscessos causam danos à pele do animal e levam à condenação da carne, principalmente quando o produto é destinado ao comércio exterior. Por isso, além do impacto econômico, a doença tem um grave impacto social (Oliveira, 2007).

Entretanto, estudos epidemiológicos recentes mostram que existe alta incidência da doença também na região sudeste. A prevalência da LC entre ovelhas no estado de São Paulo chega a 71%, enquanto a prevalência em Minas Gerais, estado onde a ovinocultura tem

apresentado grande crescimento, é de aproximadamente 75,8% entre ovinos (Guimarães-Tese 2009) e 78,9% entre caprinos (Seyffert et al., 2010).

C. pseudotuberculosis é um patógeno que pode infectar diferentes espécies animais incluindo o homem, embora seja um evento relativamente raro. Cerca de 25 casos já foram descritos na literatura e a maioria deles está relacionada ao contato direto com animais doentes ou à exposição a produtos contaminados. Liu et al (2005) relataram uma infecção por *C. pseudotuberculosis* durante um implante ocular em um paciente. Uma criança de 12 anos da França apresentou linfadenite necrosante na virilha causada por *C. pseudotuberculosis*, que foi tratada de maneira errônea pela própria falta de conhecimento do microrganismo e da patologia (Join-Lambert et al 2009).

1.2.3. Transmissão

A contaminação ambiental é considerada um fator de grande importância na disseminação da enfermidade já que *C. pseudotuberculosis* é capaz de sobreviver no ambiente por longos períodos. Sob baixas temperaturas e condições de umidade, o tempo de sobrevivência pode ser prolongado. O microrganismo pode sobreviver em frestas de piso à temperatura ambiente por até 10 dias e cerca de 8 meses ou mais de um ano em fômites (principalmente em baixas temperaturas ambiente) (Collett et al., 1994).

O microrganismo já foi isolado de alimentos, cercas, tesouras, canivetes, além de terra ao redor das áreas de manipulação dos animais (Smith e Sherman, 1994). Em bovinos e bubalinos, há evidências da transmissão da bactéria por mosca doméstica e por outros dípteros (Yeruham et al., 1996).

A principal fonte de infecção é o conteúdo dos abscessos que supuram e contaminam o ambiente. A via de entrada no hospedeiro são as feridas superficiais na pele ou as mucosas, além dos linfonodos e/ou vasos linfáticos (Pugh, 2002).

1.2.4. Tratamento e Profilaxia

Apesar de testes *in vitro* demonstrarem a sensibilidade de *C. pseudotuberculosis* a vários antibióticos como relatado no item 1.1.2, o uso destas drogas geralmente não é eficaz porque elas não penetram na rígida cápsula do abscesso e porque a bactéria localiza-se intracelularmente durante este estágio de infecção. Portanto não se recomenda o tratamento com antibióticos para os casos de LC (Collett et al., 1994). Além do que o tratamento com antibiótico é dispendioso e prolongado.

O tratamento convencional da LC é feito pela drenagem cirúrgica e posterior cauterização química com tintura de iodo. O tratamento deve ser realizado fora das instalações, o material retirado do abscesso deve ser queimado e os instrumentos desinfetados; os animais só devem retornar ao rebanho após completa cicatrização (Alves e Pinheiro, 2000).

A proposta de extirpação cirúrgica dos abscessos superficiais garantiu a cura de 90% dos animais submetidos ao procedimento. Segundo resultados, esta técnica não favorece a reincidência, além de ser mais rápida e de menor custo operacional que a drenagem cirúrgica. A prática de extirpação cirúrgica dos abscessos da LC foi utilizada por Nozaki et al (2000) em cabras visando unicamente o controle da contaminação ambiental. A cirurgia obteve êxito em 80% das cabras, entretanto os autores recomendam esta cirurgia somente quando são poucos animais acometidos no rebanho e quando a cirurgia pode ser realizada com o abscesso num estágio adequado de desenvolvimento, não sendo viável quando linfonodos e órgãos internos estão acometidos (Alves e Pinheiro, 1997).

O controle da LC deve ser realizado por medidas profiláticas com a identificação dos animais infectados, impedindo que os mesmos tenham contato com os animais saudáveis (Dorella et al., 2006a; Williamson, 2001). A primeira medida de controle baseia-se em procedimentos de manejo para evitar a disseminação bacteriana, e isso inclui a higienização de todos os instrumentos e ambientes (galpões, troncos de tosquia e baias) de uso comunitário com soluções desinfetantes após contato com os animais (Alves e Pinheiro, 1997). A tosquia deve ser iniciada pelos animais mais jovens, evitando contato entre eles e os demais, tendo o cuidado de não manipular os que possuem lesões palpáveis com os aparentemente saudáveis.

Os animais com feridas na pele devem ser banhados com água contendo agentes bactericidas. Na aquisição dos animais, deve-se levar em consideração o histórico livre da patologia e a realização da quarentena dos mesmos (Airello et al., 2001). Algumas medidas drásticas são sugeridas, como o descarte de animais que apresentem lesões aparentes ou que sejam soropositivos em testes diagnósticos. Porém, esses testes podem gerar resultados falso-positivos e consequentemente o descarte de animais não infectados (Menzies et al., 2004). Diante dos fatos mencionados, o controle deve basear-se em estratégias que impeçam a entrada e disseminação do microrganismo nos rebanhos, e para isso são necessárias medidas preventivas eficazes, como vacinas e testes diagnósticos. Apesar de intensas pesquisas laboratoriais da LC, ainda não existe uma imunoprofilaxia ideal utilizada no combate desta enfermidade (Piontkowski e Shivvers, 1998; Paton et al., 2003; Dorella et al., 2006a).

1.2.5 Diagnóstico

O diagnóstico definitivo da LC superficial é baseado na cultura bacteriológica seguida por identificação bioquímica de isolados, o qual consiste no padrão-ouro. O teste bioquímico bem estabelecido para identificação de bactérias corineformes é o sistema API Coryne (APIbioMérieux, Inc., La Balme lês Grottes, France), o qual consiste em uma bateria de 21 testes bioquímicos que podem ser realizados entre 24 e 48 horas. Assim, o desenvolvimento de um método mais rápido e mais específico para detectar *C. pseudotuberculosis* pode ser de grande valor no diagnóstico da LC.

Cetinkaya et al. (2002) propuseram um teste baseado na reação em cadeia da polimerase (PCR) para identificar isolados de *C. pseudotuberculosis* através da amplificação parcial da sequência do gene que codifica o RNA ribossômico 16S. No entanto, o teste apresentou limitações, incluindo a incapacidade de distinguir a bactéria *C. ulcerans* e a dependência de cultivo bacteriano prévio.

A amplificação de múltiplos *loci* em uma mesma reação de PCR tem sido uma estratégia amplamente utilizada para a detecção de patógenos, por permitir a identificação altamente específica do microrganismo de interesse com grande economia de tempo e de reagentes. Assim, nosso grupo de pesquisa na UFMG desenvolveu uma nova metodologia, baseada em uma reação de PCR multiplex (mPCR), para amplificar simultaneamente, sequências de três genes específicos de *C. pseudotuberculosis*: *16S rDNA* (RNA ribossômico 16S), que é o gene de escolha na maioria dos estudos de taxonomia bacteriana; *rpoB* (subunidade beta da RNA polimerase), um gene que vem sendo amplamente utilizado em análises filogenéticas de microrganismos dos gêneros *Mycobacterium* e *Corynebacterium*; e *pld* (fosfolipase D), que codifica uma exotoxina associada com a virulência em *C. pseudotuberculosis*, *C. ulcerans* e *Arcanobacterium haemolyticum*. Através da utilização deste teste de mPCR associado a um método de extração de DNA bacteriano diretamente do material purulento, foi possível detectar *C. pseudotuberculosis* mais rapidamente e de modo tão específico quanto a cultura bacteriológica seguida por identificação bioquímica dos isolados, que é considerado pelo Centro Nacional de Referência em Corynebacterioses, localizado no Instituto Pasteur, o padrão-ouro para o diagnóstico da LC (Pacheco *et al.*, 2007). Contudo, este método não é capaz de identificar o microrganismo em amostras sorológicas.

Nos últimos anos, diversos testes sorológicos foram desenvolvidos na tentativa de superar o problema do diagnóstico subclínico da LC, incluindo diferentes ensaios de ELISA, como ELISAs indiretos utilizando várias preparações da célula bacteriana, proteínas secretadas

e a toxina de *C. pseudotuberculosis* (Binns *et al.*, 2007), ELISA sanduíche com PLD (Dercksen *et al.*, 2000) e ELISA para detectar interferon gama (IFN- γ) como um marcador de imunidade mediada por células contra *C. pseudotuberculosis* (Menzies *et al.*, 2004). Contudo, a maioria destes testes não possui sensibilidade e especificidade suficientemente satisfatórias, e não permite a diferenciação entre os animais infectados e os vacinados.

1.2.6. Vacinas

Considerando o fato de o tratamento da LC ser de alto custo e ineficaz, a medida de melhor custo-benefício é a imunização. Vários relatos na literatura mostram diferentes estratégias testadas, como o uso de bactérias atenuadas, mortas, frações contendo antígenos da parede bacteriana, do sobrenadante de cultura bacteriana e uma mistura de componentes celulares e sobrenadante (Brogden *et al.*, 1990; Eggleton *et al.*, 1991; Ellis, 1991; Holstad, 1989; LeaMaster *et al.*, 1987). Essas preparações ofereceram certo grau de proteção contra infecções experimentais, mas os níveis de proteção e a severidade das lesões são variáveis. A maioria das vacinas licenciadas para fins comerciais é conjugada a outros patógenos, tais como *Clostridium tetani*, *C. perfringens*, *C. novyi* e *C. chauvoei*, sendo a eficácia contra a *C. pseudotuberculosis* não alcançada. Nestas vacinas, a exotoxina fosfolipase D encontra-se inativada (Piontkowski e Shivvers, 1998; Stanford *et al.*, 1998; Williamson, 2001). No Brasil, uma vacina viva atenuada de *C. pseudotuberculosis*, a linhagem 1002, foi licenciada desde 2000 e vem sendo industrializada, em versão liofilizada, pela Empresa Baiana de Desenvolvimento Agrícola. Mostrou uma eficiência de 83,33% na prevenção do aparecimento de abscessos superficiais, em caprinos vacinados aos 3 meses e acompanhados por um período de 8 meses (<http://www.ebda.ba.gov.br>).

O Laboratório Vencofarma também desenvolveu uma vacina contra a LC, a LINFOVAC, que é composta por bactérias vivas atenuadas o qual confere imunidade celular, e que, após vários testes, vem apresentando resultados satisfatórios em rebanhos de áreas endêmicas com proteção de até 97%. A vacinação deve ser feita em animais a partir de 3 meses de idade, revacinar 30 dias após e anualmente com 1 dose (www.vencofarma.com.br).

Com o intuito de obter uma vacina protetora e mais segura, diferentes estratégias vêm sendo testadas como vacinas recombinantes em caprinos e ovinos. Para isso, tem-se utilizado como antígeno vacinal a PLD, já que é um dos poucos fatores de virulência conhecidos e é a proteína mais bem caracterizada de *C. pseudotuberculosis* (Egen *et al.*, 1989; Holstad, 1989; McNamara *et al.*, 1995; Songer *et al.*, 1990). Chaplin *et al.* (1999) desenvolveram e testaram uma versão inativada da PLD apresentada em vacina de DNA na imunização de ovinos. O

resultado foi o de uma proteção parcial contra o desafio experimental de *C. pseudotuberculosis* similar àquela proporcionada pela vacina de subunidade inativada por formalina contra LC em ovinos (Muckle e Gyles, 1982; Paton et al., 1994; Stanford et al., 1998).

Neste contexto, a busca de antígenos imunodominantes de *C. pseudotuberculosis* tem sido realizada. A protease CP40 foi relatada como possível candidata ao desenvolvimento de vacinas, pois demonstrou proteção contra LC induzida experimentalmente (Walker et al., 1994). Da mesma forma, a proteína do choque térmico Hsp60 de *C. pseudotuberculosis* foi testada quanto ao seu potencial como antígeno vacinal, porém não conferiu proteção (Coelho, 2007), contudo novos estudos devem ser feitos e não se deve excluir a possibilidade desta proteína contribuir para imunização

Recentemente, nosso grupo de pesquisa obteve 34 mutantes de *C. pseudotuberculosis* por meio de mutagênese aleatória, utilizando o sistema de transposição baseado no TNFuZ (Dorella et al., 2006b). Após experimentos de imunização com um desses mutantes para o gene que codifica a suposta subunidade fimbrial (Cp09), garantiu 81% de proteção contra a infecção pela bactéria selvagem em camundongos o que demonstra ser um candidato promissor para a obtenção de uma nova vacina viva atenuada contra a LC (Dorella et al., 2009).

Um candidato potencial para vacina deve ser capaz de induzir uma proteção cruzada e ser abundante na célula. Além disso, deve estar localizada extracelularmente para ser acessível aos anticorpos, deve ser reconhecida pelo sistema imune e deve induzir uma resposta imune apropriada (Adamczyk-Poplawska et al., 2011). Como a regulação de genes relacionados a patogenicidade, ocorre grande parte por influência do ambiente, faz-se necessário desvendar os genes que são transcritos em diferentes condições ambientais e assim, contribuir com novos possíveis candidatos vacinais para estudos futuros de controle da doença.

1.3. Estresses Abióticos

As bactérias estão sujeitas a diversos tipos de estresses ambientais, como alterações na temperatura, disponibilidade de nutrientes, alteração do pH, presença de moléculas tóxicas, entre outros. Conseqüentemente, os mecanismos utilizados pela bactéria para responder e se proteger de diferentes tipos de estresse são essenciais para sua sobrevivência e proliferação. Genes cujos produtos estão relacionados com patogenicidade geralmente não são expressos constitutivamente, mas possuem regulação ambiental. Na maioria dos casos, a resposta bacteriana leva à ativação de genes cujos produtos respondem a um dado estresse físico-

químico. Fatores que regulam esses genes respondem a sinais específicos, ambientais (externos) ou celulares, estimulando ou inibindo a transcrição, tradução ou algum outro evento na expressão gênica de modo a modificar adequadamente a taxa de síntese dos produtos gênicos, obtendo a adaptação fisiológica e bioquímica necessária (Dow e Daniels, 2000).

A estabilidade ou meia-vida do mRNA irá influenciar nas mudanças de síntese de proteína, contribuindo para a adaptação celular e sobrevivência em resposta ao ambiente. Para a maioria das bactérias a meia-vida do mRNA está na faixa de 0,5 a 50 min, com tempo médio em cerca de 3 min (Takayama e Kjelleberg, 2000). A meia-vida é medida ou pela falha no transcrito completo (meia-vida química) ou pela perda da capacidade do mRNA servir como modelo para síntese de proteínas (meia-vida funcional) (Belasco e Brawerman, 1993).

1.3.1 Estresse osmótico

As bactérias geralmente mantêm a pressão osmótica intracelular maior do que a do meio de crescimento de forma a gerar o turgor celular necessário para diversos processos, como o crescimento e a divisão celular. O aumento da osmolaridade do meio pode desidratar o citoplasma, causando a plasmólise da célula; já em meio hipotônico, a entrada excessiva de água pode causar a lise celular. A capacidade de adaptação a mudanças na osmolaridade do meio é, portanto, importante para a sobrevivência e tem sido associada à virulência em algumas bactérias patogênicas (Sleator e Hill, 2002).

Em resposta a um aumento da osmolaridade do meio, duas estratégias têm sido descritas para a adaptação de células procarióticas: a manutenção de uma alta concentração de sal no citoplasma e o acúmulo de osmoprotetores. A primeira é utilizada por organismos de crescimento restrito a ambientes de osmolaridade elevada que toleram altíssimas concentrações de sal no meio, como as halobactérias. Já o acúmulo de solutos compatíveis (osmoprotetores) é amplamente utilizado e tem sido descrito em diversas bactérias como *E. coli*, *B. subtilis*, *Salmonella*, *Streptomyces*, dentre outras (Sleator e Hill, 2002). Solutos compatíveis são moléculas altamente solúveis, preferencialmente sem carga em pH fisiológico, que podem ser acumulados em altas concentrações sem perturbar as diversas funções celulares, auxiliando na manutenção do turgor celular (Bremer e Krämer, 2000).

A percepção da mudança na osmolaridade é mediada por proteínas osmosensoras que são proteínas integrais de membrana com extensões hidrofílicas. Alguns osmosensores atuam simultaneamente como osmoreguladores, como por exemplo, os sistemas de transporte ProP de

E. coli, envolvido no transporte de prolina, BetP de *C. glutamicum* e OpuABC de *Lactococcus lactis*, envolvidos no transporte de glicina-betaína. Estas proteínas percebem um determinado estímulo, como a concentração de íons ou propriedades da membrana, e também atuam como transportadores na tomada efetiva de solutos compatíveis do meio (Koide, 2006).

Existem também osmosensores que regulam a expressão de genes que codificam osmoreguladores, constituindo sistemas de dois componentes: o sensor localizado na membrana tem um domínio de histidina quinase que, na presença do estímulo, transmite a informação via fosforilação para os reguladores de resposta. Alguns exemplos de sistemas de dois componentes envolvidos na resposta ao estresse osmótico são: o sistema KdpDE de *E. coli* envolvido na aquisição de K^+ e o sistema EnvZ-OmpR que regula a expressão das porinas OmpC e OmpF em *E. coli*, que por sua vez, facilitam a difusão de moléculas hidrofílicas. Em resposta a um aumento na pressão osmótica, a expressão de OmpF é diminuída e OmpC tem sua expressão aumentada (Heermann e Jung, 2004). As bactérias patogênicas desenvolveram com sucesso sistemas de transdução de sinal altamente sofisticados que, em resposta aos estresses ambientais, controlam a expressão coordenada de vários determinantes de virulência. Os objetivos do patógeno são simples: sobreviver e multiplicar dentro do hospedeiro (John e John, 2002).

As mudanças na osmolaridade são uma realidade para os microrganismos que colonizam todos os ambientes e as bactérias devem ser capazes de responder imediatamente a choque osmótico, que pode ocorrer durante transições entre ambientes de crescimento, e sustentar o crescimento sob circunstâncias duradouras do estresse osmótico. A osmolaridade em ambiente aquoso equivalente a 0,06 M NaCl enquanto que no lúmen intestinal a osmolaridade é cerca de 0,3 M NaCl e na corrente sanguínea cerca de 0,15 M NaCl, em humanos. O aumento da osmolaridade está associado com expressão de fatores de virulência em muitos organismos patogênicos (Chowdhury, Sahu, Das, 1996).

Em *Salmonella flexineri*, a expressão dos genes *vir* que são necessários para invasão das células epiteliais é acentuadamente aumentada sob condições de alta osmolaridade (0,15 M NaCl) (Bernardini, Fontaine, Sansonetti, 1990).

Em um trabalho com *Pseudomonas aeruginosa* na regulação do estresse osmótico em estado estacionário, a maioria dos genes está associada com a expressão de determinantes de virulência, codificando proteínas do sistema de secreção do tipo III (TTSS), citotoxina ExoT e ExoY do tipo III e duas chaperonas (Aspedon, Palmer, Whiteley, 2006). O sistema de secreção do Tipo III é encontrado em bactérias Gram-negativas, incluindo patógenos animais,

fitopatógenos ou microrganismos simbiotes e associativos (Pühler et al., 2004) e permite a exportação de proteínas bacterianas para o interior da célula eucariótica. No interior dessas células, as proteínas bacterianas, também chamadas de efetoras, modificam sinais celulares em benefício da bactéria (Dedecek, 2006).

Genes expressos diferencialmente em resposta ao estresse osmótico podem exercer um papel chave na prevalência e na persistência da *P. aeruginosa* em locais de estresse hiperosmótico na infecção (Govan e Deretic, 1996). Na similaridade com outras bactérias, *P. aeruginosa* mostrou acumular K^+ , glutamato, trehalose, e amido de N-acetilglutaminilglutamina (NAGGN) como osmoprotetores citoplasmático em resposta ao estresse osmótico (D'Souza-Ault, Smith, Smith, 1993). Outro componente importante é a betaína, um osmoprotetor poderoso que se acumula em concentração relativamente alta no citoplasma de células sob estresse osmótico. No entanto não exerce um controle transcricional dos genes osmoprotetores. O papel dos osmoprotetores em conferir a resistência ao estresse osmótico está caracterizado relativamente bem (Csonka, 1989). Uma categoria menos conhecida de compostos osmoprotetores é a hidrofilina, uma família das proteínas básicas relativamente pequenas encontradas nas plantas, *Saccharomyces cerevisiae*, e bactérias que são induzidas pelo estresse hiperosmótico (Garay-Arroyo et al., 2000).

Um estudo do perfil transcricional de *Bacillus subtilis* crescidos no meio sob alta salinidade mostrou drástica redução na transcrição de genes *hag* que codifica uma proteína estrutural do flagelo envolvida na motilidade e quimiotaxia causando severa redução na capacidade da célula de *B. subtilis* em responder e sobreviver nestas condições. As bactérias são capazes de perceber seu ambiente e responder aos nutrientes específicos ou materiais tóxicos e mover-se em direção aos primeiros e evitar aqueles últimos (quimiotaxia positiva e negativa, respectivamente). Em um experimento com meio altamente salino e limitante de ferro, foram observadas indução de genes. Um exemplo representativo é o gene *feuA* que codifica uma proteína de transporte FeuABC que provavelmente está envolvida na absorção de ferro (Steil et al., 2003).

Em *E. coli*, células submetidas ao estresse osmótico apresentaram alterações na respiração provavelmente pela redução na transcrição de genes codificando ATP-sintase (*atpC atpF atpH* e *atpI*) e apresentaram também redução da transcrição de genes que codificam componentes da subunidade ribossomal 30S e 50S, comprometendo toda a rota metabólica (Weber e Jung, 2002).

1.3.2. Estresse por acidez

O pH citoplasmático ácido afeta negativamente o metabolismo e causa danos às macromoléculas. Muitas bactérias conseguem se adaptar ao fenômeno de queda do pH, são as bactérias tolerantes à acidez induzida (Kazmierczak, 2003). Depois da fagocitose das bactérias patogênicas pelos macrófagos, o fagossomo torna-se rapidamente ácido (onde o pH está em torno de 4,6-5,0). Apesar disto, *Mycobacterium*, por exemplo, consegue sobreviver e se multiplicar neste ambiente (Fisher, Plikaytis, Shinnick, 2002).

A resposta bacteriana envolve um processo de homeostase do pH, através do qual as células mantêm um pH intracelular relativamente constante em relação aos valores externos (Schumann e 2001). A regulação do pH citoplasmático implica em uma combinação do sistema de homeostase do pH ativo e passivo. Este sistema resulta na indução de reguladores transcricionais em resposta ao estresse ácido, exportação de prótons, mudança na composição de ácido graxo do envelope celular, síntese de proteínas de reparo, chaperonas, e produção de compostos básicos (Satorhelyi, 2004).

A exportação ativa de íons H^+ é um dos principais mecanismos para controlar o pH citoplasmático. Uma bactéria ATP-sintase tem um papel fundamental na manutenção da homeostase do pH. Em bactérias que apresentam cadeia respiratória o papel primário do complexo enzimático ATP-sintase é de sintetizar ATP a partir de gradiente de próton formado pela cadeia respiratória. Em bactéria sem cadeia respiratória, este complexo enzimático é responsável por manter a homeostase do pH via exportação do próton através da hidrólise de ATP (Martín-Galiano et al., 2001).

O sistema descarboxilase diminui a concentração de íons H^+ e regula o pH citoplasmático através da reação de descarboxilação consumindo próton. A arginina, glutamato, lisina descarboxilases usam substrato intracelular consumindo próton de cada reação. Em um pequeno grupo de bactéria a glutamato descarboxilase (GAD) ácido resistente tem sido considerada um dos principais sistemas de sobrevivência no ambiente ácido, como em *E. coli*, *Listeria monocytogenes*, *Shigella flexneri*, *Clostridium perfringens*, *Bacteroides* ssp., *Fusobacterium* ssp. e *Eubacterium* ssp. que atravessam a barreira ácida do estômago para invadir ou colonizar o intestino humano. Esta enzima mostrou atividade ótima em baixo pH, sugerindo exercer sua função em ambiente ácido (Satorhelyi, 2004).

Outro fator importante que contribui na sobrevivência da bactéria em ambiente ácido é alteração na membrana celular, resultando em redução da permeabilidade das células aos

prótons, fornecendo proteção adicional contra pH ácido (Quivey et al., 2000). Células adaptadas à acidez de *Streptococcus mutans* mostraram ser enriquecidas em mono-insaturados e longas cadeias de ácido graxo comparadas às células não adaptadas. As células de *Streptococcus* tratadas com cerulenina, um inibidor da biossíntese de ácido graxo, foram incapazes de alterar o perfil de ácido graxo da membrana e conseqüentemente incapazes de sobreviverem à severa acidificação (Fozo e Quivey, 2004). Um ensaio de permeabilidade de prótons revelou que membranas celulares mais adaptadas à acidez, são menos permeáveis aos prótons (Satorhelyi, 2004).

Existem, em bactérias, fatores de transcrição conhecidos como fatores sigma alternativo que ajudam a regular a expressão de genes em resposta a uma variedade de condições de estresse (Hecker et al., 1996). O Sigma B é um fator sigma muito comum em resposta ao estresse. Seu papel na resposta à tolerância à acidez foi estabelecido em *B. subtilis*, *Brevibacterium flavum*, *L. monocytogenes* e *Staphylococcus aureus* (Ferreira et al., 2003). Em *B. subtilis* o *SigB* regula a maioria das respostas aos estresses, contribuindo para a transcrição de mais de 100 genes. A regulação do sigma B ocorre através da atividade da proteína Rsb (regulador do sigma B). Ambientes estressantes como baixo pH são detectados pela RsbU (Vijay et al., 2000).

A rota da urease e arginina deiminase (ADI) são as rotas mais comuns em que as bactérias tentam alcalinizar seu ambiente. A neutralização dos ácidos resulta da produção de NH_3 , que combina com próton do citoplasma para produzir NH_4^+ , elevando assim o pH interno. A rota de ADI foi identificada em várias bactérias incluindo *Bacillus* spp., *Lactococcus lactis*, *Streptococcus* spp e *Pseudomonas* spp. (Satorhelyi, 2004).

A infecção pelo *Helicobacter pylori* induz inflamação persistente na mucosa gástrica com diferentes lesões orgânicas em humanos, tais como gastrite crônica, úlcera péptica e câncer gástrico. A resistência ao ácido clorídrico é de vital importância na patogênese do *H. pylori*, visto que, sem este atributo biológico, a bactéria não teria condições de colonizar a mucosa gástrica (Ladeira et al., 2003).

A enzima urease, que é uma proteína de alto peso molecular (500 a 600KDa), atua promovendo a hidrólise da uréia, presente em condições fisiológicas no suco gástrico, levando à produção de amônia. Esta atua como receptor de íons H^+ , gerando pH neutro no interior da bactéria, o que confere ao *H. pylori* resistência à acidez gástrica. Desta maneira, a bactéria fica protegida dos efeitos deletérios do pH ácido do estômago. A maior parte da urease sintetizada pela bactéria situa-se em seu citoplasma (Weeks e Sachs, 2001).

A produção de amônia depende da entrada de uréia na bactéria, que é controlada por uma proteína de membrana sensível ao pH. Esta proteína é codificada por um gene da família urease, conhecido como *ureI* (Weeks e Sachs, 2001). Cepas do *H. pylori* com deleção de *ureI* não sobrevivem em pH ácido. Weeks et al. (2000) demonstraram que a entrada de uréia na bactéria é acelerada em pH 5 e diminuída em pH 7. A entrada de uréia é altamente específica, não sendo facilmente saturada, e independe de temperatura e energia. A proteína ureI atua como portão de um canal, que também permite o refluxo de urease, aumentando o pH periplasmático e do microambiente próximo, prevenindo acúmulo tóxico de uréia dentro da bactéria.

A bactéria, na fase precoce de colonização, necessita atravessar a camada de muco que protege o epitélio gástrico. Tal camada é formada por um gel visco elástico que confere proteção química e mecânica ao revestimento epitelial, inclusive contra bactérias (Jenks e Kusters, 2000). No entanto lipases e proteases sintetizadas pelo *H. pylori* degradam a camada de muco, facilitando a progressão da bactéria.

A colonização da mucosa gástrica pelo *H. pylori* requer um complexo processo adaptativo. O fator-chave, que permite à bactéria sobreviver à acidez gástrica e atravessar o lúmen do estômago, é a enzima urease, como já dito, que converte a uréia em amônia e bicarbonato (Tombola et al., 2001). Para isso é necessária a produção da citotoxina VacA, que induz a formação de canais seletivos de ânions nas células epiteliais, levando à exsudação de uréia para a luz da mucosa gástrica. A VacA é considerada importante fator de virulência, visto que contribui para a produção de alcalóides pela urease, que podem induzir danos no DNA (Tombola et al., 2001).

Saviola et al (2003) trabalhando com *Micobacterium tuberculosis* foram capazes de isolar duas supostas regiões do promotor as quais levaram a indução de genes que foram especificamente expressos em meio ácido, sendo silenciados em pH 7,0, mas induzidos em pH 4,5. Os genes foram identificados como *lipF* e *rv0834c*. O gene *lipF* está anotado no genoma de *M. tuberculosis* como lipase ou esterase. Como *lipF* foi induzido em pH ácido, uma possível função seria de hidrolisar os ácidos graxos tóxicos presentes em necrose caseosa durante a tuberculose; alternativamente pode modificar a parede celular externa como uma resposta adaptativa aos prejuízos ácidos. O gene *rv0834c* está anotado no genoma de *M. tuberculosis* como um gene de PE-PGRS que pode estar envolvido na modulação do sistema imune, porém necessita de mais estudos.

O sistema de reparo do DNA e as moléculas chaperoninas foram relacionados à resistência ácida de bactérias (Hanna et al., 2001). A acidificação intracelular aumenta a

depurinação do DNA. A depurinação manifesta-se em ruptura da cadeia e em ligações cruzadas, causando uma perda local de informação genética. Nas bactérias sob condições de estresse ácido, vários mecanismos de reparo do DNA são potencialmente envolvidos no reparo do DNA induzido pela acidez, incluindo reparo por recombinação (RecA), reparo via excisão de base (BER), reparo do pareamento alterado das bases (MMR- mismatch repair), reparo via excisão de nucleotídeo (NER), que tem a capacidade de localizar e retirar as lesões do DNA (Sancar, 1996).

1.3.3. Estresse térmico

Temperatura é um dos parâmetros físicos sob constante vigilância em todas as células. Numerosos processos são controlados por temperatura nas bactérias incluindo a expressão de genes do choque térmico, choque pelo frio e genes de virulência, para mencionar os exemplos mais proeminentes. Três moléculas diferentes foram identificadas como sensores térmicos: DNA, mRNA e proteínas (repressores transcricional, chaperonas e protease) (Narberhaus et al., 2006).

A resposta ao choque térmico é induzida por aumentos da temperatura. Quanto mais elevada a temperatura, mais pronunciado é o nível da indução dos genes de choque térmico. Outra característica é que a resposta ao choque térmico é transitória e breve (Schumann, 2007). Um exemplo é quando as células de *E. coli* são transferidas de 30°C a 42°C, ou células de *B. subtilis* de 37°C à 48°C. O crescimento da *C.pseudotuberculosis* é observado na faixa de temperatura entre 14 e 40°C, porém seu crescimento ótimo se dá em uma temperatura de 37°C (Alves et al., 1997) e a temperatura corporal dos caprinos e ovinos pode variar de 37 a 40°C (Santos, 2006).

Em muitos patógenos, os determinantes de virulência estão sob o controle dos ativadores transcricional que respondem às flutuações de temperatura crescente que conduz a uma expressão elevada de genes de virulência a 37°C. Em *L. monocytogens*, um patógeno intracelular facultativo Gram positivo, a ativação de genes de virulência está sob o controle do ativador transcricional PrfA (Leimeister-Wächter et al., 1992). O gene que codifica PrfA é transcrito através do seu próprio promotor como um transcrito monocistrônico em 30°C. Em 37°C, além do transcrito monocistrônico, o gene *prfA* é transcrito também a partir de um promotor diferente como uma mensagem bicistrônica que compreende genes *prf* e *pic*

codificando proteínas (fosfolipase C fosfatidilinositol-específico) que atuam rompendo a membrana do hospedeiro através hidrólise dos lipídeos. Logo, o gene do *prfA* é transcrito a partir de duas regiões diferentes do promotor em situações de altas temperaturas e o transcrito bicistrônico que é termicamente regulado.

Os mecanismos alternativos para a ativação dos fatores da transcrição de genes de virulência foram descritos em *Y. enterocolitica* e em *V. cholerae*. Em *Y. enterocolitica* o gene *virF* é o ativador transcricional dos genes de virulência. A transcrição deste gene é auto-regulada pela temperatura (Cornelis et al., 1989). Há evidências que sugerem que o aumento relativamente abundante da transcrição do gene *virF* devido a temperatura crescente, pode ser devido às alterações da super hélice do DNA (Dorman, 1991). Em contraste, em outros organismos patogênicos como em *V. cholera*, *Y. pseudotuberculosis* e *Y. enterocolitica*, os genes de virulência são otimamente expressos em 30°C, mas reprimidos a 37°C (Parsot e Mekalanos, 1990; Pierson e Falkow, 1990).

Efeito Choque Térmico

O efeito choque térmico é uma resposta celular protetora e homeostática em oposição aos estresses que levam a prejuízos conformacionais de proteínas. Ocorre quando células que crescem em temperaturas baixas são transferidas para um ambiente de alta temperatura e resultam na indução de um grupo de genes denominado genes do choque térmico codificando proteínas do choque térmico (HSPs) (Schumann, 2007).

As chaperonas têm por função proteger outras proteínas na obtenção de sua conformação apropriada. A estrutura primária de HSPs parece ser altamente conservada durante a evolução, sugerindo que elas apresentam funções similares em todos os organismos (Morimoto, 1998). Muitas são chaperonas (DnaK, GroEL, proteína Clp) ou proteases dependente de ATP (Lon, FtsH, ClpAP, ClpXP) e exercem papel importante no enovelamento de proteína, agrupamento, transporte, reparo, tanto em condições de estresse quanto sem estresse (Schumann, 2007). GroEL junto com GroES facilita translocação da proteína atravessando barreiras na membrana e possivelmente também secreção. As funções fundamentais destas HSPs são: impedir a desnaturação da proteína e evitar proteínas desnaturadas parcialmente reativa. As proteínas desnaturadas não reparadas são degradadas por outra classe de HSPs representada ou por protease dependente do ATP (ex. Lon, La) ou uma subunidade catalítica (ClpP, Ti) ou regulatória (ATPase) (ClpB) ou outra protease Clp (Craig et al., 1993).

A resposta ao choque térmico tem sido estudada extensivamente, tanto em bactérias Gram-negativas como Gram-positivas. Em *E. coli*, a expressão dos genes das Hsps é regulada positivamente pelo fator sigma alternativo σ_{32} , codificado pelo gene *rpoH*. A regulação dos níveis de σ_{32} é exercida por uma retroinibição onde as Hsps (DnaK, DnaJ e GrpE) direcionam σ_{32} para proteólise mediada por FtsH sob condições normais de temperatura (Gross, 1996). Como os níveis de DnaK e DnaJ são limitantes na célula, o aumento dos níveis de proteínas desenoveladas durante o choque térmico sequestra essas chaperonas, liberando o σ_{32} para ligar-se à RNA polimerase, direcionando-a para promotores específicos de genes de choque térmico (Yura e Nakahigashi, 1999). É possível que em altas temperaturas o fator σ_{32} livre, tenha alta afinidade pela RNA polimerase e a clivagem por FtsH seja impedida, levando ao aumento da meia-vida do fator σ_{32} para cerca de 4 - 5 minutos a 42°C que em condições normais, a uma temperatura de 30°C, é cerca de 1 minuto (Bukau, 1993). Além disso, o aumento da temperatura provoca um rápido aumento na tradução de σ_{32} devido à desestabilização da estrutura secundária do mRNA de *rpoH*, liberando o acesso dos ribossomos ao sítio de início da tradução do mensageiro (Morita et al., 1999).

Quando os níveis de DnaK aumentam, o fator σ_{32} é novamente sequestrado, sendo inativado e degradado, desligando a resposta. Foi demonstrado que a maquinaria das chaperoninas GroES/EL também está envolvida na regulação da atividade e estabilidade de σ_{32} em *E. coli* (Guisbert et al., 2004). Porém, a dependência por fatores sigmas na regulação de choque térmico não é conservada em procariotos. Em diversos organismos, muitos genes *hsp* importantes são controlados exclusivamente por repressores específicos, como em *Bacillus* e *Streptomyces* (Servant e Mazodier, 2001).

Proteínas do choque térmico são proteínas codificadas em resposta a elevação de temperatura ou em outra situação crítica celular. A razão para este comportamento é que o enovelamento da proteína é afetado pelo calor e, portanto, algumas chaperonas agem reparando o dano potencial causado pela falha no enovelamento. Outras chaperonas estão envolvidas no enovelamento de novas proteínas que deixam o ribossomo. Apesar da maioria das novas proteínas poderem definir sua conformação na ausência de chaperonas, uma estrita minoria as requer. O efeito choque térmico é uma resposta ao aumento da temperatura, que é transitória e resulta na expressão de genes que permite as células removerem proteínas desnaturadas (Fenton e Horwich, 2003).

1.4. Projeto Genoma de *Corynebacterium pseudotuberculosis*

A pesquisa genômica vem crescendo cada vez mais no cenário mundial. Com o surgimento de novas tecnologias, os bancos de dados genéticos vêm sofrendo incrementos nunca antes imaginados. Atualmente, no Brasil, existem diversos centros de pesquisa utilizando a tecnologia de nova geração (*Next-Generation Sequencing*, ou *NGS*) como: Laboratório Nacional de Computação Científica (Petrópolis/RJ) – Sistema 454; Universidade Federal do Pará (Belem/PA) - plataforma SOLiD™; Centro de Pesquisas René Rachou (BH/MG) – plataforma SOLiD™; Instituto Ludwig (São Paulo/SP) - plataforma SOLiD™; Instituto Israelita de Ensino e Pesquisa Albert Einstein (São Paulo/SP) – plataforma SOLiD™; Instituto de Biologia Molecular (Curitiba/PR) – plataforma SOLiD™ que possivelmente deixarão o país em uma posição ainda mais privilegiada. Estas tecnologias apresentam uma habilidade de processar milhões de sequências em uma única corrida (Mardis, 2008a)

No Brasil, duas das Redes Genomas mais atuantes são: a Rede Paraense de Genômica e Proteômica (RPGP), apoiada pela FAPESPA e a Rede Genoma de Minas Gerais (RGMG) apoiada pela FAPEMIG.

Em 2006, a RGMG apoiou o sequenciamento do genoma da bactéria *Corynebacterium pseudotuberculosis*, linhagem 1002, na tentativa de caracterizar geneticamente o agente etiológico da LC, com a perspectiva de identificar proteínas relacionadas à patologia e que pudessem contribuir para o desenvolvimento de uma eventual vacina, de terapias e de kits de diagnóstico.

O projeto foi finalizado em maio de 2009 e o genoma da linhagem 1002 de *C. pseudotuberculosis* já se encontra depositado no banco de dados do NCBI sob o número de acesso CP001809.

Interessante ressaltar que, apesar da sua inegável importância na sanidade animal, essa bactéria é pouco caracterizada e poucos genes envolvidos nos mecanismos patogênicos são realmente conhecidos. Por isso a importância de se empreender tentativas para caracterizar geneticamente o microrganismo e contribuir para o controle da doença ou mesmo erradicação.

Existem hoje, 14 espécies pertencentes ao gênero *Corynebacterium* que possuem os seus genomas completamente sequenciados depositados no NCBI. Em 2007 eram apenas quatro, o que nos mostra que o interesse em caracterizar diferentes espécies do gênero vem crescendo (<http://www.ncbi.nlm.nih.gov>). No caso da espécie *C.pseudotuberculosis* só alguns genes/proteínas encontram-se depositados. O nosso grupo LGCM possui hoje diferentes linhas de pesquisa envolvendo *C. pseudotuberculosis*, utilizando diferentes abordagens- genômica,

pangenômica, transcriptômica, proteômica, e possivelmente será o maior e principal grupo, a contribuir com os bancos de dados público, armazenando novas informações a respeito deste organismo.

Em maio de 2009, foi concluído o primeiro sequenciamento de nova geração na América Latina, utilizando o sistema SOLiD™ (*Supported Oligonucleotide Ligation and Detection*) da Applied Biosystems, que é uma das plataformas de sequenciamento genômico que utilizam tecnologia NGS. Foram decodificados naquele momento, os genomas de duas novas linhagens de *C. pseudotuberculosis*, isoladas de cavalo (linhagem 258) e camelo (linhagem 162) e em parceria com o LGCM estes organismos foram montados, anotados e curados em prazo recorde de 21 dias.

Graças a esta parceria, além da linhagem 1002, já foi depositada também no NCBI a linhagem CpC231 sob o número de acesso CP001829, a linhagem I19 sob número de acesso CP002251, a linhagem PAT10 sob o número de acesso CP002924, linhagem CIP5297 sob número de acesso CP003061. Os outros genomas de *C. pseudotuberculosis* estão sendo finalizados e brevemente estarão disponíveis no NCBI. O quadro 1 informa todas as linhagens que foram sequenciadas (exceto C231,106, 399 e 4202), montadas e anotadas até o momento, pelo grupo LGCM em parceria com RGMG.

QUADRO 01: Linhagens que foram sequenciadas (exceto C231, 106, 399 e 4202), montadas e anotadas até o momento pelo grupo LGCM em parceria com RGMG

Espécie	Linhagem	Biovar	Isolado	Local	Origem	Metodologia
<i>C. pseudotuberculosis</i>	1002	tipo I	Caprino	Abcesso	Bahia	Sanger/454
<i>C. pseudotuberculosis</i>	C231	tipo I	Ovino	desconhecido	Austrália	454
<i>C. pseudotuberculosis</i>	162	tipo II	Camelo	Abscesso	Reino Unido	SOLiD/Fragmento
<i>C. pseudotuberculosis</i>	258	tipo II	Equino	desconhecido	Bélgica	SOLiD/Fragmento
<i>C. pseudotuberculosis</i>	CIP5297	tipo II	Equino	desconhecido	Quênia	SOLiD/Pareada
<i>C. pseudotuberculosis</i>	PAT10	tipo I	Ovino	desconhecido	Patagônia	SOLiD/Pareada
<i>C. pseudotuberculosis</i>	I19	tipo I	Bovino	Mastite	Israel	SOLiD/Pareada
<i>C. pseudotuberculosis</i>	31	tipo II	Bufalo	desconhecido	Egito	SOLiD/Fragmento
<i>C. pseudotuberculosis</i>	31	tipo II	Bufalo	Abscesso	Egito	Ion / Fragmento
<i>C. pseudotuberculosis</i>	267	tipo I	Lhama	desconhecido	Califórnia	SOLiD/Fragmento

<i>C. pseudotuberculosis</i>	316	-	Cavalo	desconhecido	California	Ion / Fragmento
<i>C. pseudotuberculosis</i>	106	tipo II	Cavalo	Abscesso	Escocia	454/Illumina
<i>C. pseudotuberculosis</i>	399	tipo I	Caprino	Abscesso	Escocia	454/Illumina
<i>C. pseudotuberculosis</i>	4202	tipo I	Ovino	Abscesso	Escocia	454/Illumina

1.5. Plataforma SOLiD™

A plataforma SOLiD™, lançada em 2007 pela Applied Biosystems (<http://www.appliedbiosystems.com>) difere das demais por realizar a leitura da sequência de nucleotídeos de uma fita de DNA durante a reação de incorporação de dinucleotídeos marcados, catalisada pela DNA ligase. A incorporação de um dinucleotídeo é seguida da excitação do fluoróforo, da leitura do sinal e da remoção do fluoróforo antes da incorporação do dinucleotídeo seguinte. O resultado da leitura dos sinais fluorescentes gera um código de cores que é analisado na forma de uma matriz de cores para então ser transformado no tradicional código de letras (Mardis, 2008b). O sistema SOLiD™ é o único desta geração que tem a capacidade de distinguir entre um erro no sequenciamento e um verdadeiro polimorfismo de nucleotídeo simples (SNP) devido a vantagem de interrogar cada base duas vezes em ciclos independentes de iniciadores fornecendo confiabilidade nos resultados. Ele usa quatro cores fluorescentes para codificar 16 combinações possíveis entre duas bases (Figura 01).

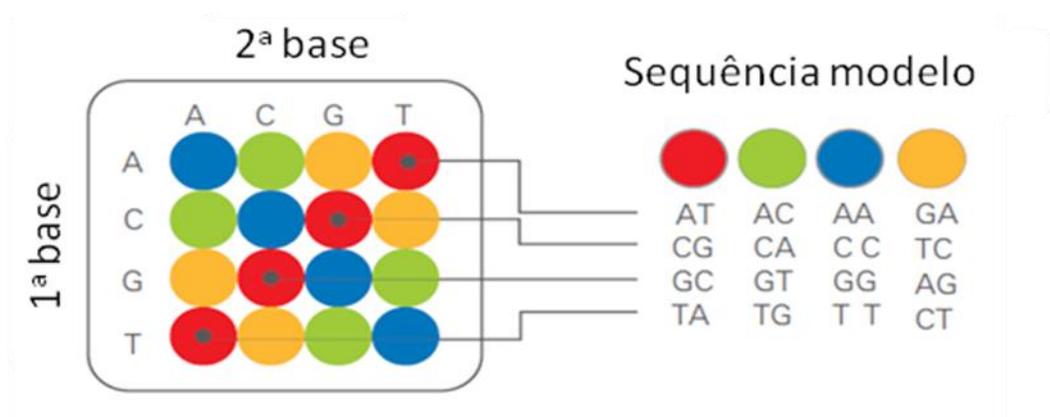


FIGURA 1 - Quatro cores utilizadas no sistema SOLiD™ para detectar as duas bases, em 16 possíveis combinações.

A versão 2 do SOLiD™ é capaz de gerar de 3 a 5 Gb (GigaBase) e leituras de cerca de 35 bases de extensão, a um custo aproximado de 2 mil dólares por corrida. O sistema SOLiD™

foi premiado, em 2008, como o melhor na categoria de instrumentação para análise genômica (<http://www.lifescienceindustryawards.com/2008/winners.html>).

Através da versão 3 Plus do SOLiD™, foi possível elevar ainda mais o desempenho do sequenciamento, tanto no sentido de reduzir o tempo de corrida, de 8 para 6 dias, mas também no sentido de gerar leituras maiores (~70 bases) e com um rendimento que pode chegar a 100 Gb (<http://www.appliedbiosystems.com>).

1.5.1. Uso do SOLiD™ para a análise do transcriptoma

Uma aplicação revolucionária do SOLiD™ é a análise de transcriptoma, com uma eficiência comparável a dos métodos tradicionalmente utilizados, como PCR em tempo real (RT-PCR). Este sistema fornece aos pesquisadores o método revolucionário com alta sensibilidade para caracterizar todo o transcriptoma de um organismo, descobrir novos transcritos, identificar mutação, deleção e inserção, splicing alternativo, com excelente cobertura gerando mais de 600 milhões de leituras em uma única corrida (www.appliedbiosystems.com). Transcriptoma é o conjunto de todos os transcritos da célula (codificantes ou não-codificantes), e sua quantificação, para um estágio de desenvolvimento específico ou condição fisiológica. Entender o transcriptoma é essencial para interpretar os elementos funcionais do genoma e revelar os constituintes moleculares da célula, e assim, compreender o desenvolvimento e as doenças (Wang, Z. et al., 2009).

Uma vez que o sequenciamento por si só não nos informa sobre a função ou papel do gene diante de determinadas condições, o próximo passo seria a investigação baseada no processo pós-genômico, ou seja, estudar a expressão dos genes codificados pelo genoma do patógeno em determinadas condições fisiológicas, tentando entender a resposta a essas condições.

Nas últimas décadas, as técnicas para avaliação de expressão gênica avançaram bastante no que diz respeito ao volume de dados obtidos que se expande de um ou poucos genes, por exemplo, através de Northern blotting, RT-PCR quantitativa, Ensaio de Proteção a Nuclease até obtenção de um grande número de genes analisados, através de Hibridização Subtrativa, Diferencial Display, SAGE, Microarranjo (Moody, 2001). Contudo, estamos nos aproximando dos limites técnicos da tecnologia do microarranjo e estes estão sendo substituídos por transcriptômica baseada em tecnologia do sequenciamento de DNA de nova geração em “ultra larga escala”.

O termo RNA-seq tem sido utilizado para representar o transcriptoma desvendado por sequenciamento da nova geração (*NGS*). Embora pioneiro em organismos eucariotos devido à maior facilidade de isolar o mRNA destes organismos, a tecnologia de sequenciamento de RNA já vem sendo aplicada em procariotos (Van Vliet, 2010).

A metodologia baseada em hibridização (microarranjo), considerada larga-escala é relativamente barata, com exceção dos arranjos de alta resolução (*tiling-arrays*) que investigam grandes genomas. Porém, apresenta limitações: alto índice de ruído devido à hibridização cruzada, limitação na faixa de detecção devido ao próprio ruído e saturação de sinais, não detecta os transcritos com baixo número de cópias por célula, detecta apenas os transcritos representados nos suportes e não garante cobertura total dos transcritos. Além disso, as análises de expressão comparativa entre diferentes experimentos são difíceis e requerem complicado método de normalização baseados em cálculos estatísticos muito complexos (Van Vliet, 2010).

Embora seja recente o estabelecimento da nova geração de sequenciamento, ela já demonstra uma extraordinária preferência em todas as áreas da pesquisa genômica e tudo indica que substituirá a tecnologia dos microarranjos (Teng e Xiao, 2009), devido as vantagens que a tecnologia de RNAseq proporciona, como: ter pouco ou quase nenhum ruído e oferecer a possibilidade de mapeamento inequívoco das sequências em única região do genoma. Garante também alta cobertura e grande faixa de detecção de transcritos, podendo detectar desde uma cópia de RNA por célula até grande número de cópias por célula, a custos consideravelmente mais baixos em comparação aos métodos tradicionais (Mane et al., 2009). Simplifica muito o processo de preparação das amostras pela remoção da etapa de clonagem minimizando erros (Morozova e Marra, 2008). Além disso, o RNAseq tem se mostrado altamente preciso na quantificação absoluta dos níveis de transcritos, obtendo resultados similares à detecção através da PCR quantitativa (qPCR) (Wang et al., 2009). A RT-PCR quantitativa era a única tecnologia referência para avaliar os níveis de transcrição, apesar de não ser um ensaio perfeito, e para (Roberts et al., 2011), é a melhor opção além do próprio RNAseq. Hoje, RNAseq já vem sendo considerada padrão-ouro para análise dos níveis de expressão (Sendler et al., 2011).

Para gerar o mapa do transcriptoma, as leituras originadas no sequenciamento de RNAs são mapeadas no genoma de referência e as regiões expressas são determinadas através da cobertura destas leituras. O método ideal para análise transcriptômica deve ser capaz de identificar diretamente e quantitativamente todos os RNAs (codificantes e não-codificantes), raros e abundantes, menores e maiores e a nova geração de sequenciamento de RNA permite

estas análises. Por todas as vantagens oferecidas, o RNAseq é o primeiro método que permite examinar todo o transcriptoma de maneira rápida, quantitativa e com precisão, em custo muito menor do que os arranjos ou sequenciamento de EST em larga escala por Sanger (Wang, Z. et al., 2009).

Hoje, a análise de transcriptoma em toda a extensão do genoma tem sido conduzida por duas técnicas, principalmente: *tiling-arrays* ou RNA-seq (Sorek e Cossart, 2010).

A tecnologia do *tiling-array*, um subtipo do microarranjo, que representa ambas as fitas do genoma em alta densidade, já foi explorada no estudo de transcriptoma de *L. monocytogenes* (Toledo-Arana et al., 2009), *B. subtilis* (Rasmussen et al., 2009), *Halobacterium salinarum* (Koide et al., 2009), *Mycoplasma pneumoniae* (Güell et al., 2009) fornecendo excelentes e importantes resultados. No entanto, a pesquisa baseada nos arranjos necessita de centenas de milhares de sondas e é limitada por ruídos, hibridização cruzada e requer extensa normalização. A vantagem é que não precisa enriquecer os mRNAs, como no RNAseq. No entanto, a maioria dos *tiling-arrays* tem baixa densidade, principalmente devido ao custo associado ao grande número de sondas necessárias. Por isso, o mapa de transcriptoma gerado pela técnica de *tiling-array*, pode ser de baixa resolução comparada à utilização do RNA-seq (Sorek e Cossart, 2010).

Toledo-Arana et al (2009) analisaram a transcrição do genoma inteiro de *L. monocytogenes* utilizando a tecnologia do *tiling-array* e descobriram uma diversidade de RNAs: pequenos RNAs envolvidos na virulência, RNAs antisense em sobreposição com ORFs (janelas abertas de leitura) e sobreposição com 5' e 3' UTR (regiões não traduzidas). Ainda neste trabalho foi descoberto que "*riboswitches*", que são elementos regulatórios do mRNA que controlam a expressão de grande número de genes em resposta a vários metabólitos celulares (Serganov, 2009), podem atuar como terminadores dos genes e várias outras características e funções que vão ajudar a elucidar os mecanismos de regulação em bactérias. Estes pesquisadores conseguiram demonstrar a reformulação da transcrição no processo de infecção e a relevância do sigma B no controle da expressão de genes de virulência importantes para *L. monocytogenes* na sua adaptação ao ambiente intestinal. Ao contrário, no sangue, a proteína PrfA quem controla a transcrição dos genes de virulência envolvidos na sobrevivência e replicação neste ambiente. Surpreendentemente, vários RNAs não-codificantes ausentes em uma espécie não patogênica, *L. innocua*, exibiram o mesmo padrão de expressão dos genes de virulência. Estes dados desvendam as alterações transcricionais que ocorrem durante um processo de infecção em uma bactéria, até então desconhecidos (Toledo-Arana et al., 2009).

Uma descoberta interessante, através do estudo de transcriptoma, foi obtida a partir do genoma de *M. pneumoniae*. Güell et al (2009) utilizaram duas técnicas, tiling arrays e sequenciamento direto de fita-específica (DSSS), para analisar a bactéria. Pela “hibridação” das técnicas conseguiram observar a expressão de todos os genes. A análise revelou a versatilidade de operons em resposta a diferentes condições (foram testadas 173 condições diferentes), ou seja- um gene codificado como policistrônico em uma condição, pode ser transcrito como monocistrônico em outra. Esta versatilidade do operon foi observada em mais de 40% das transcrições de *M. pneumoniae* e já foi documentado em um estudo do transcriptoma em arquea (Koide et al., 2009). Estes relatos reforçam a visão dos operons como estruturas alternativas (não estáticas) que aumentam a capacidade regulatória do transcriptoma de modo que é funcionalmente análogo aos promotores alternativos ou de *splicing* alternativo em transcriptoma de eucariotos. A disponibilidade de ferramentas experimentais baseadas em transcriptoma para determinar operons aumentará ainda mais o entendimento da regulação versátil e da evolução de transcritos policistrônico em procariotos.

Em *B. subtilis*, Rasmussen et al (2009) encontraram 3662 regiões transcricionalmente ativas e confirmaram 77,3% dos genes anotados no genoma, revelando 84 supostos RNAs não codificantes e 127 transcritos antisenso.

A eficiência da análise do transcriptoma no refino da anotação de genomas indica que esta ferramenta se tornará componente padrão no procedimento de anotação de genomas. A maioria dos genomas de procarioto sequenciados é anotada quase que exclusivamente por preditores gênicos, processo em que o software pode identificar genes codificadores de proteínas, um pequeno grupo de RNAs não codificantes (ncRNA) e pequenos RNAs (sRNA). Estas anotações estão sujeitas a erros, não conseguem detectar todos os pequenos genes e deixam de identificar a maioria dos ncRNAs e regiões UTRs (Sorek e Cossart, 2010).

Embora o campo da transcriptômica esteja nos seus primórdios, já está claro que fornece meios para compreensão de mecanismos regulatórios baseados no RNA. O estudo do transcriptoma nos aproxima bem da realidade pela descoberta de novos genes, por definir a estrutura do gene anotado, corrigindo a primeira metionina, detecção de UTRs, incluindo *riboswitches* e sitio ligante de sRNAs regulatórios e por determinar a ligação dos operons envolvidos na fisiologia ou patogenicidade de qualquer procarioto. A tecnologia de RNAseq ingressa na pesquisa no intuito de garantir resultados ainda melhores, mais próximos da realidade e por um custo menor que a tecnologia do *tiling array* (Sorek e Cossart, 2010).

Até recentemente, acreditava-se serem extremamente raros os casos de transcrito cis-antisense em procariotos. Os casos raros dos transcritos cis-antisense avaliados, revelaram a participação em importantes processos da regulação, tais como: replicação, resposta ao estresse e transporte de ferro (Brantl, 2007). Graças à revolucionária análise de transcriptoma total, utilizando as novas tecnologias em ultra larga escala, já se sabe da abundância destes transcritos em bactérias e arqueas. Um RNA transcrito a partir da fita senso pode interagir com o RNA antisense, e pode ser uma forma comum de regulação para procariotos (Sorek e Cossart, 2010).

Entre as descobertas utilizando a tecnologia de *NGS*, destacam-se os pequenos RNAs (small RNAs ou sRNAs) que mostraram claramente seu papel na fisiologia dos procariotos. Poucos relatos existiam a respeito de sRNAs em procariotos pela dificuldade de detecção. Estes sRNAs que apresentam um tamanho entre 50 e 500 pares de bases, regulam importantes processos biológicos, tais como virulência, resposta ao estresse e “*quorum sensing*” (Bejerano-Sagie e Xavier, 2007; Toledo-Arana et al., 2007). Os relatos afirmam que funcionam de forma análoga aos miRNAs (micro RNAs) em eucariotos e, hoje, já se sabe da abundância destes elementos em organismos procariotos (Sorek e Cossart, 2010).

A abordagem da tecnologia *RNA-seq* permitiu encontrar em *Salmonella typhi* 25 genes que apresentavam uma extremidade 5' UTR mais longa- normalmente a região 5' UTR em bactérias são menores do que 30 pares de bases- e dois destes genes estavam presentes em regiões genômicas de alta concentração de genes de virulência, o que sugere um papel desta região UTR na regulação da virulência (Perkins et al., 2009).

Embora em procariotos, a atenção esteja voltada principalmente para estudos envolvendo 5'UTR ou sitio inicial da transcrição (TSS), tudo indica que futuramente, os estudos baseados em transcriptoma revelarão também o papel regulatório da extremidade 3' UTR (Sorek e Cossart, 2010). Uma extensa região 3' UTR pode afetar a expressão de genes que estejam localizados na fita oposta (Toledo-Arana et al., 2009). Em arqueas, transcritos com região 3' UTR de tamanho considerável revelaram ter um papel na regulação da tradução (Koide et al., 2009; Wurtzel et al., 2010).

Oliver et al (2009) evidenciaram como a metodologia do RNAseq pode fornecer excelentes resultados na caracterização quantitativa do transcriptoma de procarioto, fornecendo novas estratégias para explorar a rede regulatória transcricional em bactérias. Eles utilizaram a tecnologia Illumina Genome Analyzer para caracterizar o transcriptoma de *L. monocytogenes* 10403S na fase estacionária e de outra linhagem desta espécie desprovida de fator σ^B , que está

relacionado a mecanismos gerais de resposta a estresses. Neste trabalho, os autores conseguiram demonstrar que 83% de todos os genes foram transcritos na fase estacionária e que 42% dos genes anotados no genoma mostraram um nível de transcrição que foi de médio a elevado nestas condições. Um total de 96 genes tiveram níveis significativamente elevados de transcrição em relação ao mutante, indicando a transcrição dependente do sigmaB destes genes. As análises do RNAseq indicaram um total de 67 moléculas de RNA não-codificantes, incluindo 7 supostamente não conhecidas. Os dados do RNAseq foram capazes também de melhorar a anotação de prováveis operons, bem como a visualização de regiões 5'- e 3'- UTR.

Outra contribuição importante utilizando a estratégia de RNA seq, através de Illumina Genome Analyzer e SOLiD™, foi o trabalho de (Passalacqua et al., 2009). Este foi o primeiro trabalho disponível a comunidade científica, utilizando o sistema SOLiD. Eles conseguiram demonstrar o transcriptoma total de *Bacillus anthracis* sob uma variedade condições de crescimento e forneceram com acurácia e alta resolução o mapa de transcritos únicos e estrutura de operons através do genoma. Além disso, identificaram sequências previamente não anotadas no genoma, que se apresentaram com significativa atividade transcricional, como 37 regiões não traduzidas na extremidade 5' (5'UTR) maiores que 100 pares de base, que podem ser indicativas da capacidade funcional nesta região, garantindo maior qualidade na anotação do genoma da espécie; demonstraram também, a quantificação absoluta dos transcritos, e sugeriram que existe uma heterogeneidade transcricional significativa dentro de um clone.

Isabella e Clark (2011) conseguiram detectar um grande número de genes diferencialmente expressos no genoma de *Neisseria gonorrhoeae*, crescendo em ambiente anaeróbico, condição que é incomum para este organismo. A sensibilidade da tecnologia de sequenciamento de RNA, pelo sistema SOLiD™, permitiu um número de leituras de 1.490.403 em média, entre quatro experimentos, e assim, revelou reguladores transcricionais envolvidos no crescimento anaeróbico, confirmou a expressão e regulação de pequenos RNAs e muitos genes envolvidos na adaptação e resposta ao estresse anaeróbico, além de muitas proteínas hipotéticas induzidas nesta condição, podendo assim serem alvos de futuros estudos. Este trabalho permitiu também a identificação de genes induzidos, que foram expressos a partir do plasmídeo pJD1 de *N. gonorrhoeae*.

Estudos com os transcritos da bactéria Gram-negativa *Acinetobacter baumannii*, trouxeram novos resultados para a ciência. Os autores testaram o efeito do etanol no perfil transcricional pelo sequenciamento de RNA através do Illumina. Obtiveram 3.596.474 leituras mapeadas unicamente no genoma o que permitiu detectar a expressão de 49 genes induzidos

pertencentes a diferentes categorias funcionais, e 21 genes reprimidos. Comprovaram que o etanol é assimilado eficientemente e influencia na virulência e crescimento da bactéria. Os resultados confirmaram que a aplicação do RNAseq é uma poderosa ferramenta para identificar os principais moduladores envolvidos na patogenicidade bacteriana e que mesmo ocorrendo uma contaminação das amostras com grande número rRNA, ainda que tenha sido filtrado, as leituras que alinham unicamente garantem a profundidade do sequenciamento e qualidade dos resultados (Camarena et al., 2010)

Uma inovação no RNAseq foi realizada com o patógeno *Helicobacter pylori* (Sharma et al., 2010), onde foram construídas duas bibliotecas: uma tratada com exonuclease, que elimina moléculas com 5' monofosfato (rRNA e tRNA) e portanto enriquecida com transcritos primários que apresentam a extremidade 5' trifosfato, e uma outra biblioteca com RNA total sem tratamento com exonuclease, conseguindo assim, representar o mapa de resolução de nucleotídeos deste patógeno humano através do sequenciamento pela Roche FLX. Aproximadamente 217 milhões de leituras do cDNA foram mapeadas no genoma. Para contribuir na predição de operons, sequenciaram também no Illumina amostras de cDNA obtidas a partir de RNAs fragmentados aleatoriamente. Nesta pesquisa encontraram grande número de sitio inicial de transcrição (TSS) nunca antes imaginado. Observaram que a região 5'UTR, variando da TSS até o start códon apresenta, em sua maioria, um tamanho entre 20 a 40 nucleotídeos e encontraram correlação entre tamanho da 5'UTR com a função celular. Tamanho maior está relacionado à patogenicidade, por exemplo. Outras descobertas interessantes garantiram uma nova visão da organização do transcriptoma de *H. pylori* e forneceram dados para uma melhor análise de genes individuais. Foi revelado que o transcriptoma desta bactéria é altamente complexo apesar do genoma ser pequeno e compacto.

Por todas as vantagens relatadas aqui, apresentamos o estudo da genômica funcional de *C. pseudotuberculosis* 1002, biovar tipo I (ovis), pelo sistema SOLiD™. Porém, neste trabalho será apresentado apenas o perfil transcricional diferencial de *C. pseudotuberculosis* 1002. Assim, detectamos o “*estimulon*” que é o conjunto de genes ou operons que são diferencialmente expressos em resposta a uma perturbação ambiental (Zhou e Yang, 2006). Além disto, determinamos também o *core stimulon*, que definimos como o conjunto de genes comuns entre as três condições de estresses. Definimos o padrão de expressão destes genes, buscando elucidar o processo de infecção, desvendar a modulação da resposta bacteriana frente às alterações ambientais sofridas durante o processo de instalação da doença, verificando seu escape do sistema imune no hospedeiro e, assim, contribuir também para o desenvolvimento de

projetos futuros como desenvolvimento de vacinas mais eficazes e novas terapias, além de permitir a correção e atualização da anotação no genoma, em trabalhos posteriores.

2. JUSTIFICATIVA

A LC é reconhecida como uma das principais causas de perdas econômicas significativas na caprinovinocultura mundial (Baird e Fontaine, 2007). O tratamento da enfermidade baseado na antibioticoterapia não é eficiente; sendo o controle através de medidas de manuseio e profilático os mais adequados para combater a patogenia (Baird e Fontaine, 2007; Williamson, 2001).

As vacinas contra a *C. pseudotuberculosis* ainda não são totalmente eficazes na redução da incidência da doença (Dorella et al., 2009; Paton et al., 2003; Piontkowski e Shivvers, 1998). Alguns ensaios de diagnósticos também foram elaborados, padronizados e testados (Baird e Fontaine, 2007; Dorella et al., 2006a, Menzies et al., 2004; Pacheco et al., 2007; Williamson, 2001), porém possuem resultados divergentes e há limitação quanto à detecção da enfermidade em animais sem sintomatologia aparente. Essa inexistência de métodos profiláticos para o controle da LC compromete o manejo eficaz dos animais e contribui para a disseminação da doença.

Pouco se sabe como os determinantes de virulência de *C. pseudotuberculosis* são regulados ou sobre o papel destes no processo da doença. O fato é que estas bactérias patogênicas conseguem se adaptar e sobreviver no ambiente intracelular escapando da resposta imune do hospedeiro contra a infecção instalada.

Sendo assim, nosso grupo de pesquisa tem se dedicado a estudos cada vez mais aprofundados deste microrganismo. A proposta deste trabalho foi caracterizar a expressão gênica diferencial de *C. pseudotuberculosis* linhagem 1002 (biovar tipo I) em algumas condições que simulam o ambiente encontrado pela bactéria no hospedeiro no processo de infecção (estresse osmótico, térmico e acidez), e assim investigá-lo e revelar a modulação da resposta para sobrevivência nestas condições. Este estudo pode ainda contribuir com os estudos futuros relacionados ao controle ou erradicação da doença por revelar novos possíveis determinantes de virulência.

Para se atingir estes objetivos era necessário e indispensável concluir a primeira etapa do projeto que foi sequenciar, montar e anotar o genoma envolvido. Partimos primeiramente para finalização da primeira etapa - a Genômica e somente após concluirmos esta, iniciamos a segunda etapa - a Transcriptômica. Os resultados da Genômica estão apresentados nos anexos de 1 a 5.

3. OBJETIVO

Caracterizar os genes diferencialmente expressos de *C. pseudotuberculosis* em condição fisiológica de cultivo e sob diferentes condições ambientais.

3.1. Objetivos específicos

- Avaliar a replicação da linhagem 1002 de *C. pseudotuberculosis* em diferentes condições de estresse ambiental: choque térmico, osmótico, acidez.
- Extrair o RNA total e enriquecer o mRNA
- Sintetizar e amplificar o cDNA para sequenciamento no SOLiD™
- Descrever o *core estimulon* entre as três condições de estresse
- Detectar o “*estimulon*” do biovar tipo I sob as diferentes condições de estresses pré-determinados comparando com a condição fisiológica de cultivo

4. MATERIAL E MÉTODOS –Etapa transcriptoma

4.1 Linhagem utilizada

Como organismo modelo deste trabalho foi utilizada a bactéria *Corynebacterium pseudotuberculosis* linhagem 1002 (CP001809) isolada de cabras infectadas no Brasil. Utilizamos a versão 13 depositada no NCBI para o alinhamento dos transcritos.

4.2 Curva de crescimento da linhagem 1002

Partindo de uma colônia isolada semeada em BHI sólido foram preparados pré-inóculos da linhagem 1002 selvagem em 20 mL de meio BHI líquido acrescido de Tween[®] 80 à concentração final de 0,05%. Após 24 horas de crescimento em incubadora tipo shaker a 37°C, parte dos cultivos foram re-inoculados (1:100) em 100 mL de meio BHI acrescido de Tween[®] 80 (0,05% final) e incubados a 37°C. A cada hora, a cultura foi monitorada através de leituras de densidade ótica (DO_{600nm}) até atingir o início da fase estacionária, o que ocorreu 14 horas após a inoculação. Os dados de DO_{600nm} foram representados com auxílio do *software* GraphPad Prism v.5.0.

4.3 Avaliação da susceptibilidade de *C. pseudotuberculosis* aos estresses

As diferentes condições de estresse *in vitro* utilizadas para avaliar a resistência da linhagem 1002 de *C. pseudotuberculosis*, cultivada em meio BHI, foram determinadas em um trabalho realizado por Pacheco (Tese, 2010) em colaboração com a Universidade de Warwick, Reino Unido e baseado no trabalho desenvolvido por Castro (Dissertação, 2009), optou-se pela fase inicial de crescimento exponencial para a extração do RNA total, para análise do perfil transcricional em diferentes condições de estresses.

4.3.1 Experimentos de resistência ao estresse térmico

Inicialmente, um pré-inóculo foi diluído 1/100, em meio BHI fresco acrescido de Tween[®] 80 à concentração final de 0,05% e cultivado a 37 °C em incubadora tipo shaker, até atingir o início da fase exponencial de crescimento (DO_{600nm}= 0,2). Logo após atingir a fase desejada, a cultura foi dividida em diferentes frações nas quais foram aplicadas as diversas condições de estresse e em seguida, o tubo foi centrifugado, o sobrenadante descartado e o

precipitado bacteriano foi ressuspenso com o meio BHI acrescido de Tween[®] 80 à concentração final de 0,05% a 50°C e mantido nesta temperatura por 15 min na incubadora tipo shaker. Imediatamente após 15 min, foi retirada uma alíquota de 100 µL da cultura e submetida a diluições seriadas em 900 µL de caldo BHI fresco, sendo realizados plaqueamentos em duplicatas para as diluições 10⁻⁴, 10⁻⁵ e 10⁻⁶ com a finalidade de verificar a viabilidade celular. As placas de Petri semeadas foram mantidas em incubadora a 37°C até a visualização de colônias. Retirou-se ainda 12 mL da cultura, centrifugou e ao precipitado bacteriano acrescentou 6 ml de *RNAlater*[®] (Ambion, EUA), um reagente para estabilização do RNA, para posterior extração de RNA. Procedimento idêntico foi realizado para o controle (cultura não submetida ao estresse). Os dados de viabilidade foram representados com auxílio do *software* GraphPad Prism v.5.0.

4.3.2 Experimentos de resistência ao estresse osmótico

A partir da cultura distribuída (item 4.3.1), após centrifugar e descartar o sobrenadante, o precipitado bacteriano foi ressuspenso com o meio BHI acrescido de Tween[®] 80 à concentração final de 0,05% a 37°C contendo 2M de NaCl, e foi mantido na incubadora tipo shaker por 15 minutos. O procedimento então foi o mesmo relatado no item 4.3.1

4.3.3 Experimentos de resistência ao estresse por acidez

A partir da cultura distribuída (item 4.3.1), após centrifugar e descartar o sobrenadante, o precipitado bacteriano foi ressuspenso com o meio BHI acrescido de Tween[®] 80 à concentração final de 0,05% a 37°C pH 5,0 (suplementado com HCl) e mantido na incubadora tipo shaker por 15 minutos. A partir deste ponto, o procedimento foi o mesmo relatado no item 4.3.1

4.4. Obtenção do material para extração do RNA

Resumidamente, as alíquotas de 12mL foram retiradas dos tubos de 50mL, centrifugada e ao precipitado bacteriano, foram adicionados 6 mL de *RNAlater*[®] (Ambion, EUA). Após misturar em *vortex* por 5 segundos, os tubos foram incubados a 4°C *overnight*, para

estabilização da atividade transcricional e conservação da integridade do RNA. As amostras foram então submetidas à centrifugação a 12000 rpm por 5 minutos, seguida de descarte do sobrenadante. Os precipitados bacterianos estabilizados foram armazenados em freezer -80°C até o procedimento de extração do RNA total.

4.5. Extração do RNA total

As bactérias estabilizadas com *RNAlater*[®] (Ambion, EUA) foram submetidas à extração de RNA total com o emprego do *kit* ChargeSwitch[®] Total RNA Cell –Invitrogen (EUA) sendo seguidas as recomendações do fabricante com as seguintes adaptações. Após adição do tampão de lise o material foi transferido para tubos de 2 mL parcialmente preenchidos por microesferas de vidro de 1 mm de diâmetro (Bertin Technologies, FR). A lise mecânica das células foi feita no homogeneizador “Prescellys 24” configurado para agitar os tubos a 6.500 rpm por 2 ciclos de 15 s, com intervalo de 30 s entre estes. As amostras foram então centrifugadas por 1 min e o sobrenadante dos tubos transferido para novo tubo de 2 mL e incubadas em banho seco a 60° C por 15 min seguindo todo protocolo original. Quando indicado, foram acrescentadas DNase para a eliminação de DNA genômico residual. A eluição do RNA total presente nas *beads* magnéticas foi feita com 100 µL de água mili-Q RNase-Free. A quantidade de RNA total obtido foi avaliada no equipamento Qubit[®] 2.0 Fluorometer-Invitrogen (EUA)

4.6. Remoção do rRNA e obtenção do cDNA

Para produção da biblioteca de cDNA é necessário enriquecer o mRNA e assim, cerca de 2 µg do RNA total de cada amostra foi depletado do rRNA, utilizando o *kit* da Invitrogen Ribominus[™] Transcriptome Isolation (Yeast and Bacteria) (EUA) seguindo recomendações do fabricante, onde o RNA ribossomal foi suprimido através de hibridização com sondas específicas marcadas com biotina e posterior captura destes complexos por *beads* magnéticas ligadas a estreptavidina. O RNA total não contendo o rRNA foi utilizado para síntese do cDNA utilizando o protocolo padrão do *kit* SOLiD[™] Total RNA-Seq (Applied Biosystem, EUA). O material foi quantificado em Qubit[®] 2.0 Fluorometer-Invitrogen (EUA).

4.7. Preparo da biblioteca de cDNA

Na figura 02 segue um *pipeline* resumindo os passos para a obtenção da biblioteca que foi construída utilizando o protocolo *SOLiD Total RNA-Seq Kit*. Resumidamente, após a remoção dos ribossomais, cerca de 245 ng do RNA foi fragmentado com a RNase III através de incubação em um termociclador a 37°C por 10 minutos e interrompida pela adição de 90 µl de água sem nuclease. O RNA foi concentrado através da precipitação por etanol segundo *RiboMinus Concentration Module* (Invitrogen-EUA), para hibridização com adaptadores e posterior preparação de amplificação da biblioteca de cDNA, produzidos por transcrição reversa, a partir de adaptadores P1 ligados à extremidade das moléculas de RNA, conforme protocolo *SOLiD™ Total RNA-Seq Kit* (Applied Biosystem, EUA). Após esta etapa, preparou-se uma eletroforese em gel de poliacríamida desnaturante 6% para recuperação de fragmentos na faixa de 150 a 250 pares de base para amplificação do cDNA através de uma PCR. A amplificação foi realizada em um termociclador através de 15 ciclos de 95°C por 30 segundos, 62°C por 30 segundos e 72°C por 30 segundos, sucedidos por incubação a 72°C por 7 minutos. Finalmente, o material amplificado foi purificado com o kit *PureLink™ PCR Micro* (Invitrogen- EUA) e o material foi quantificado em *Qubit® 2.0 Fluorometer*-Invitrogen (EUA) e confirmado através de eletroforese de agarose 2%. As amostras foram então ajustadas para uma concentração de 60pg/mL, sugerida para a PCR em emulsão.

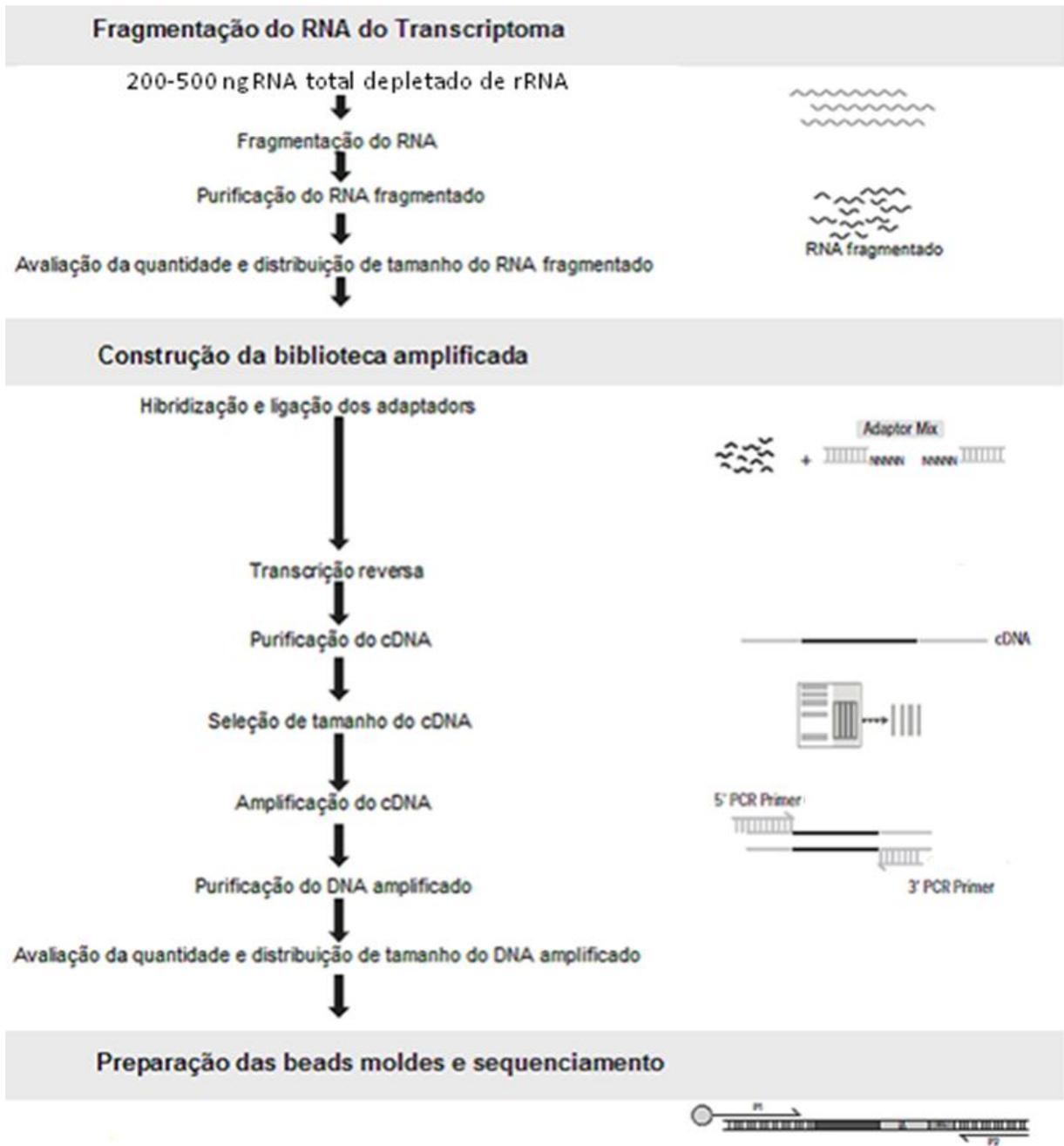


FIGURA 2- Pipeline resumindo os passos para preparo das bibliotecas no SOLiD™
 Fonte: modificado do *SOLiD Total RNA-Seq Kit Protocol*

4.8 Preparo das microesferas para o sequenciamento

Seguiu-se então para amplificação por PCR em emulsão em *full-escala* (produz de 150 a 300 milhões de microesferas de interesse), utilizando-se primers complementares aos adaptadores P1 das microesferas, conforme recomendações da *Applied Biosystems SOLiD™ 3 Plus System Templated Bead Preparation Guide*. Segue-se, na figura 3, os passos da preparação.

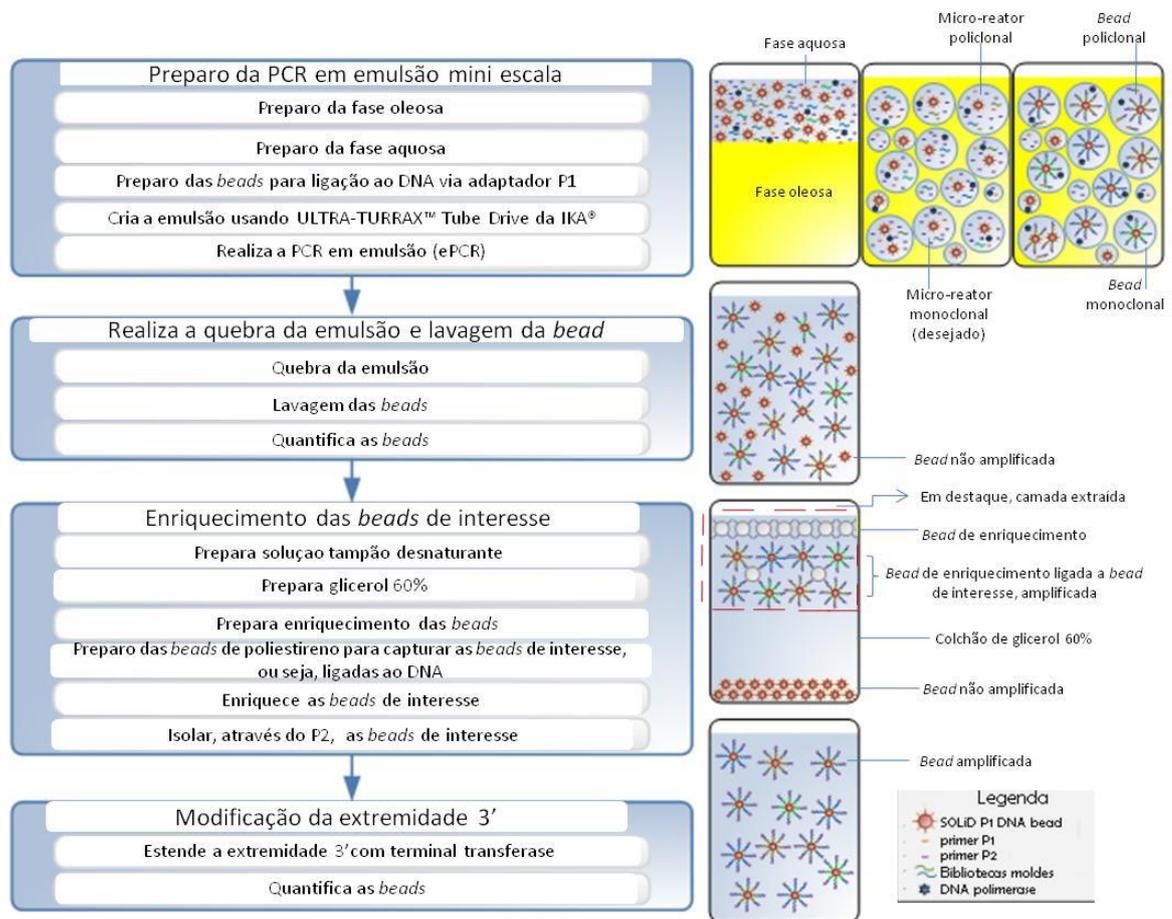


Figura 3: Descrição dos passos para preparo da PCR em emulsão

Figura obtida e adaptada do manual *Applied Biosystems SOLiD™ 3 Plus System Templated Bead Preparation Guide*

Resumidamente, a fase oleosa e aquosa são preparadas separadamente, então emulsifica utilizando o sistema ULTRA-TURRAX™ Tube Drive da IKA®, de forma a gerar microgotículas de fase aquosa no interior da emulsão (micro-reatores). Após 5 minutos de emulsificação a reação foi distribuída em placas de 96 poços e levadas ao termocilador. A amplificação foi realizada em um termociclador através de 40 ciclos de 93°C por 15 segundos,

62°C por 30 segundos e 72°C por 75 segundos, sucedidos de uma extensão final a 72°C por 7 minutos.

Ao término da reação de emulsão foi feita uma reação de enriquecimento, onde apenas as microesferas com fragmentos inseridos são capturadas. Em seguida, correu-se a WFA (*workflow analysis*) onde 15.000.000 dessas microesferas foram depositadas em lâminas e passaram por um ciclo de ligação para avaliação de quantidade e qualidade. Com base nesta avaliação, 44.000.000 de microesferas foram depositadas para o sequenciamento conforme descrição do fabricante, produzindo leituras de 50 nucleotídeos.

4.9. Análise dos dados

Os transcritos obtidos a partir do sequenciamento com SOLiD™ V3 foram analisados quanto a qualidade através do software Quality Assessment (Ramos *et al.*, 2011) e tratados no programa SAET (SOLiD™ Accuracy Enhancer Tool), uma ferramenta que aumenta a acurácia dos dados brutos de saída do SOLiD™. Este programa corrige erros (bases erradas) dentro das leituras, antes de mapeá-las no genoma e após este procedimento foram realizadas as análises *in silico* e testes estatísticos através do *software* BioScope™ (Software for Scientists Guide) descrito pelo fabricante SOLiD™ Applied Biosystems. Este software tem um conjunto de programas para análise das leituras obtidas no experimento de transcriptoma, uma ferramenta de mapeamento das leituras para criar o melhor mapa de transcritos. Resumidamente (Figura 04):

- Filtro: Muitas leituras do transcriptoma são RNA ribossomal e outras fontes de RNA sem interesse. Para manter estas leituras fora do arquivo de interesse, este programa identifica as leituras que foram filtradas. Estas não entram no arquivo final de mapeamento. O filtro é para conter as sequências sem interesse, como sequências dos adaptadores, RNAs ribossomais ou transportadores.
- Alinhamento: As leituras nesta etapa são alinhadas contra o genoma de referência.
- Contagem dos genes: Neste passo, identifica o número de leituras que alinharam nas regiões transcricionalmente ativas do genoma (quantos genes no genoma)
- Cálculo da cobertura: Reporta a cobertura das leituras em cada posição genômica

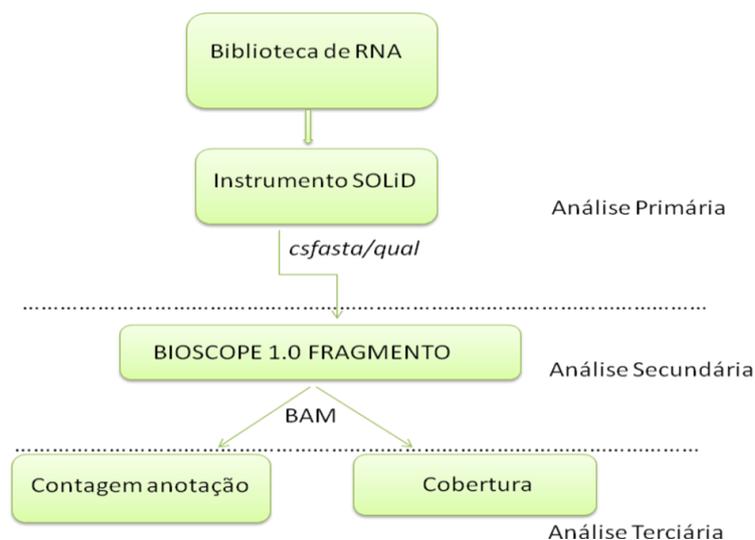


FIGURA 4- Pipeline do BioScope.

Os arquivos de saída do SOLiD (*csfasta* e *.qual*) serão arquivos de entrada para o programa BioScope, juntamente com o genoma de referência de *C.pseudotuberculosis* 1002 e o arquivo contendo as sequências dos ribossomais que foram filtradas. A análise secundária fornece informação de mapeamento, conhecido também como alinhamento das leituras. A análise terciária permite visualizar e analisar os dados, incluindo contagem das tags e determinação da cobertura

4.10. Análise da expressão diferencial e seleção dos genes diferencialmente expressos

Para identificação dos *estimulons* utilizamos o programa DEGseq (Wang et al., 2009), um pacote R para identificação de genes diferencialmente expressos a partir dos dados do RNAseq. O método requer as leituras mapeadas, a partir dos dados de RNAseq no genoma referência, ou pode fornecer o valor de expressão do gene obtido por outros programas, como RPKM (Mortazavi et al., 2008). No caso deste trabalho, utilizou-se o arquivo de saída do Bioscope com valores de RPKM, como arquivo de entrada para o DEGseq, podendo-se assim, inferir os valores da razão entre os valores de RPKM dos genes expressos na condição de estresse sob RPKM da condição controle.

Partindo do pressuposto que cada leitura é uma amostra independente, os dados podem ser modelados como processo de amostra aleatória. Assim, o número de leituras segue uma distribuição binomial e os dois modelos estatísticos de Teste Exato de Fisher e Razão de Verossimilhança estão integrados no DEGseq, contribuindo para identificação dos genes diferencialmente expressos (Wang et al., 2009).

O arquivo de saída contém, além do valor de expressão (RPKM), o valor de *p* calculado para cada gene que são confirmados pelos valores de *q* para correção de testes múltiplos que propõem uma medida de significância estatística nos estudos do genoma baseados no conceito de Taxa de Falsa Descoberta –FDR (*False Discovery Rate*) definida como a proporção

esperada de falso-positivo. Desta forma, oferece um equilíbrio entre o número de verdadeiros e falsos positivos (confiabilidade estatística) (Storey, 2002).

O usuário pode definir tanto pelo *p-value* ou *q-value* (baseado na FDR) para identificar os genes diferencialmente expressos entre as bibliotecas, utilizando um valor de corte de 0,001. Neste trabalho utilizou-se o *p-value* para detecção dos alvos, o qual mede a taxa de falso positivo que é a taxa de hipótese nula verdadeira que foi classificada como significativa. Com este valor de corte, Hedenfalk et al. (2001) demonstraram que cerca de três falsos positivos são esperados. O programa informa também o valor da razão de expressão (*fold-change*) em logaritmo de base 2 ($\log_2\text{fold-change}$) de uma condição em relação à outra.

Para o armazenamento das informações de expressão do transcriptoma, foi modelado um banco de dados utilizando o MySQL. A utilização do *DBTransProt* acelerou de forma significativa as análises dos níveis de expressão dos genes, pois além de mostrar os resultados de uma forma organizada e rápida, evita a análise manual dos arquivos de saída do programa Bioscope e DEGseq (Sá, dados não publicados).

Foi desenvolvido também o programa CoreStImulon (CSI) (Pinto, dados não publicados), que no caso deste trabalho foi utilizado para identificar genes que compõem os processos biológicos, função molecular e componente celular, obtidos através do Blast2GO (www.blast2go.org). Para tal, o programa recebe como arquivos de entrada, os transcritos de cada condição e gera relatórios analíticos por processo biológico, função molecular e componente celular. Neste trabalho, visualiza-se apenas o resultado gerado para o processo biológico. No entanto, o programa permite outras análises quando se realiza a montagem *ab initio* dos transcritos (anexo 7), como a detecção de novos transcritos no genoma. Este software foi implementado utilizando a linguagem de programação JAVA com a biblioteca Swing (<http://java.sun.com/>) e para a geração dos relatórios foi utilizado o software IReport (<http://java.sun.com/docs/books/tutorial/uiswing/>).

5. RESULTADOS

5.1. Curva de Crescimento

A curva de crescimento realizada em meio BHI com a linhagem 1002 está representada na figura 05. A linhagem *C.pseudotuberculosis* 1002 apresentou uma replicação após uma hora de inoculação, na condição controle, em concordância ao trabalho de Pacheco (Tese, 2010) e Castro (Dissertação, 2009).

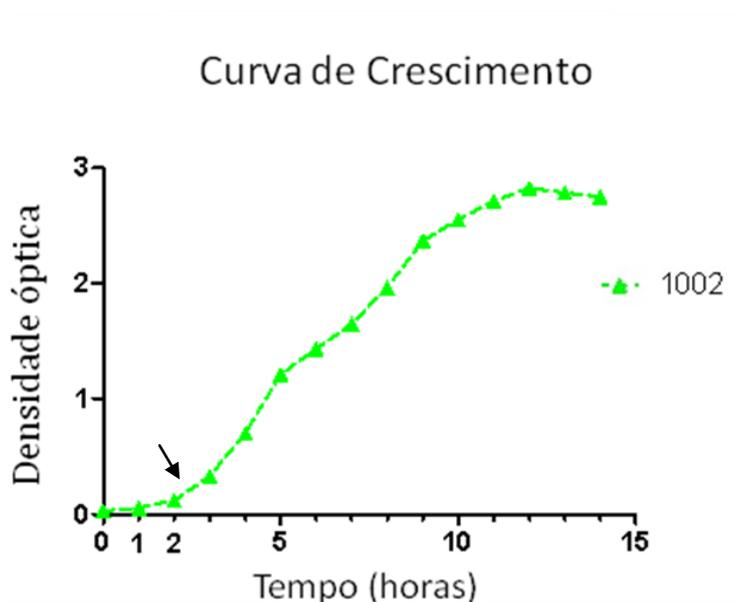


FIGURA 5- Curva de crescimento de *C.pseudotuberculosis* linhagem 1002 em condição controle, medida em densidade óptica de 600 nm (DO_{600nm}). Os triângulos indicam o valor da densidade a cada hora. A seta indica o ponto onde foram aplicados os estresses ($DO_{600nm}=0,2$); início da fase exponencial.

5.2. Avaliação da susceptibilidade de *C. pseudotuberculosis* aos estresses

Após atingir a DO ($A_{600}=0,2$), aplicou-se os agentes geradores de estresse (item 4.3.1; 4.3.2; 4.3.3). Através do experimento de viabilidade celular, podemos inferir que houve uma redução na replicação em torno de 27% para estresse térmico, 34% na condição de acidez, e 23% no estresse osmótico. Os dados podem ser visualizados na figura 6.

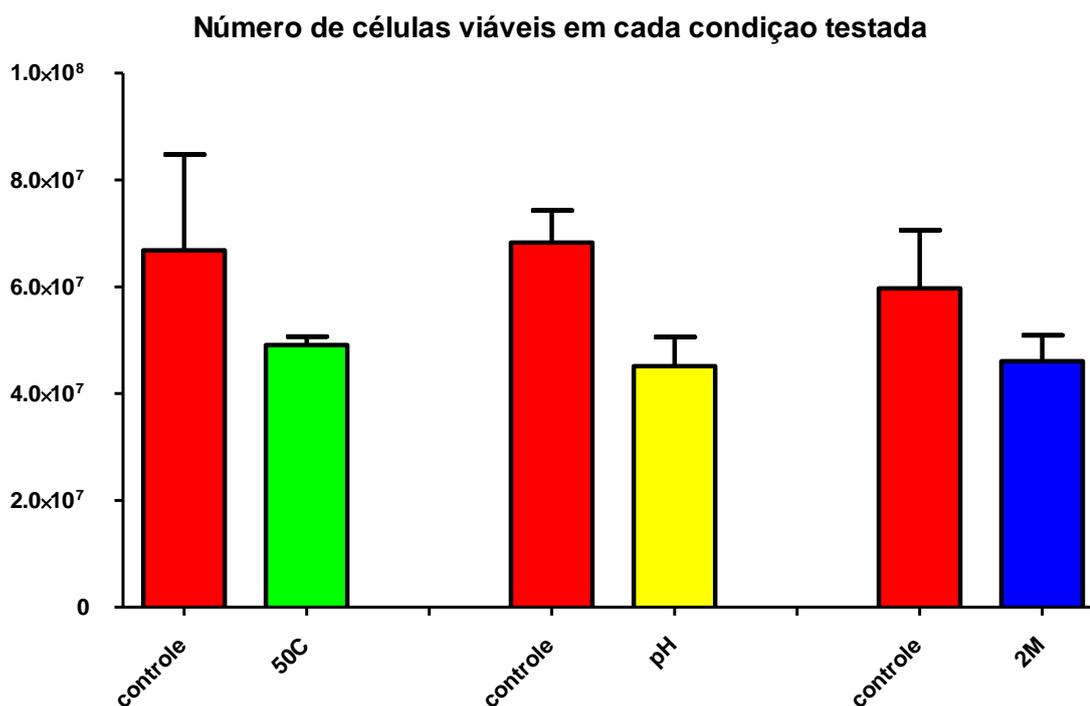


FIGURA 6- Número de células viáveis em cada condição.

Condição controle equivale, em média, a $6,70 \times 10^7 \text{ ml}^{-1}$ células da linhagem 1002, exceto controle do osmótico que equivale a $6,0 \times 10^7 \text{ ml}^{-1}$. Cor vermelha-condição controle. Verde- estresse térmico. Amarelo- estresse ácido. Azul- estresse osmótico. Figura obtida no programa GraphPad Prism (<http://www.graphpad.com/prism>).

5.3. Extração do RNA

Na figura 7 podemos observar a integridade do RNA total e remoção do DNA através do tratamento com DNase. Observa-se duas bandas relacionadas aos ribossomais (23S e 16S) e na figura 8 nos mostra o RNA depletado de RNAs ribossomais através do *Kit* da Invitrogen Ribominus™ Transcriptome Isolation (Yeast and Bacteria), o qual visualiza-se uma banda (16S) mais intensa, confirmando a redução dos ribossomais. Porém, o *kit* garante remoção de 95% dos ribossomais, mas ainda visualizamos uma quantidade razoável de ribossomais.

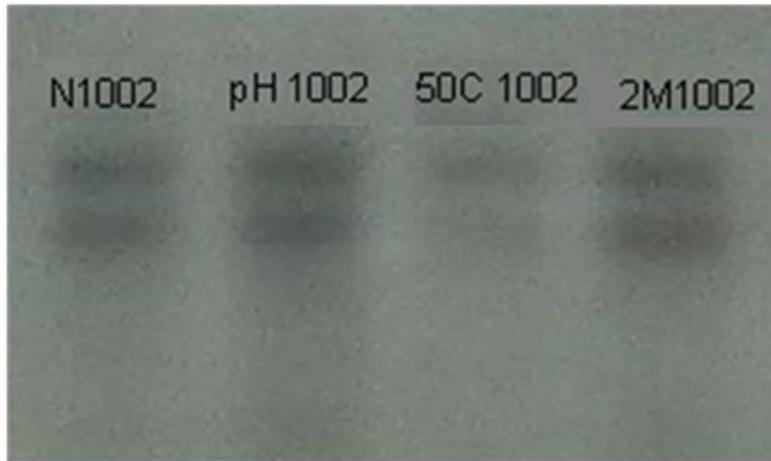


FIGURA 7 - RNA total, tratado com DNase.

N1002-amostra controle. pH 1002-amostra obtida no estresse ácido. 50C°-amostra obtida no estresse térmico. 2M-amostra obtida no estresse osmótico. Observa-se apenas as bandas referentes a 23S e 16S no gel.

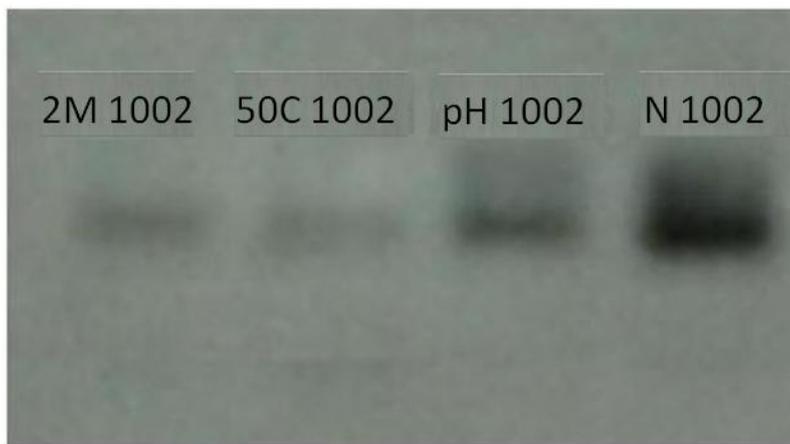


FIGURA 8 - RNA depletado dos ribossomais.

N1002-amostra controle. 50°C-amostra obtida no estresse térmico. pH 1002-amostra obtida no estresse ácido. 2M-amostra obtida no estresse osmótico. Observa-se mais intensa a banda referente ao 16S, o 23S foi mais depletado

5.4. Processamento dos dados do SOLiD™ e Análise em Softwares

Cada amostra de RNA foi sequenciada utilizando o sistema SOLiD™, e a qualidade de deposição das *beads* pode ser observada na figura 8, onde foi muito similar entre estresse térmico e controle (e talvez por isso, a cobertura entre os dois esteja próxima) e ainda se observa uma qualidade melhor de deposição das *beads* nestas condições, pois as *beads* estão localizadas na linha tracejada (figura 9), indício de *beads* monoclonais (ideal para o experimento). De posse dos dados, as leituras foram submetidas ao programa Quality Assessment (Ramos et al., 2011) (figura 10) para se observar a qualidade das leituras.

Observamos que a média de qualidade ficou na faixa de Phred 20 e a quantidade de leituras produzidas pode ser visualizada na tabela 1.

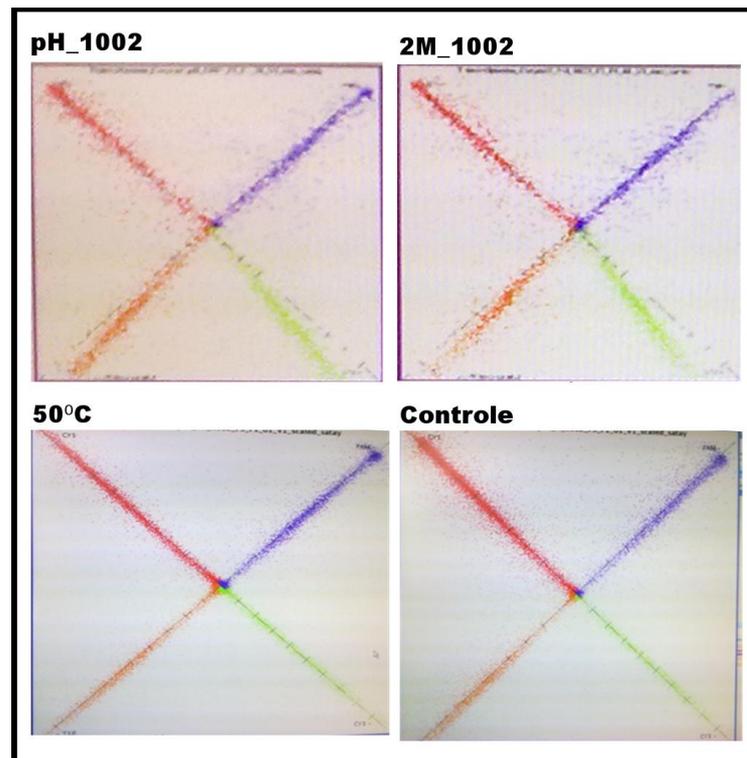
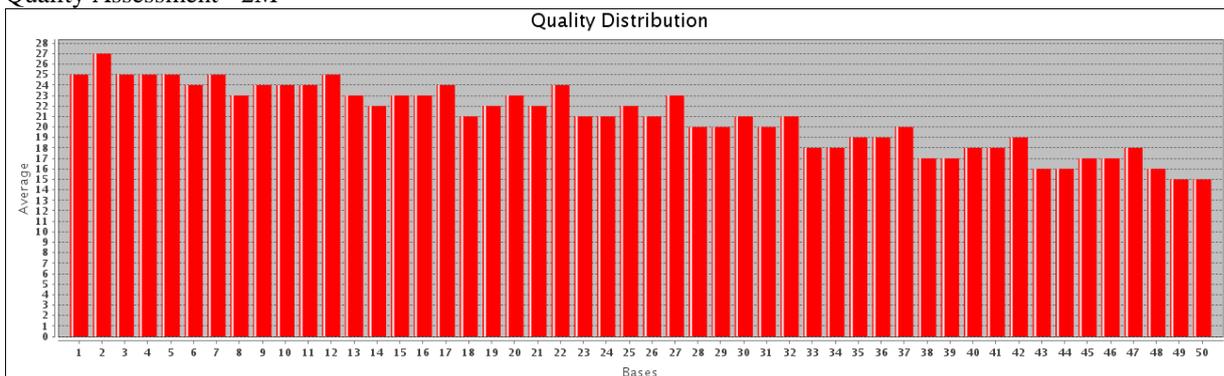
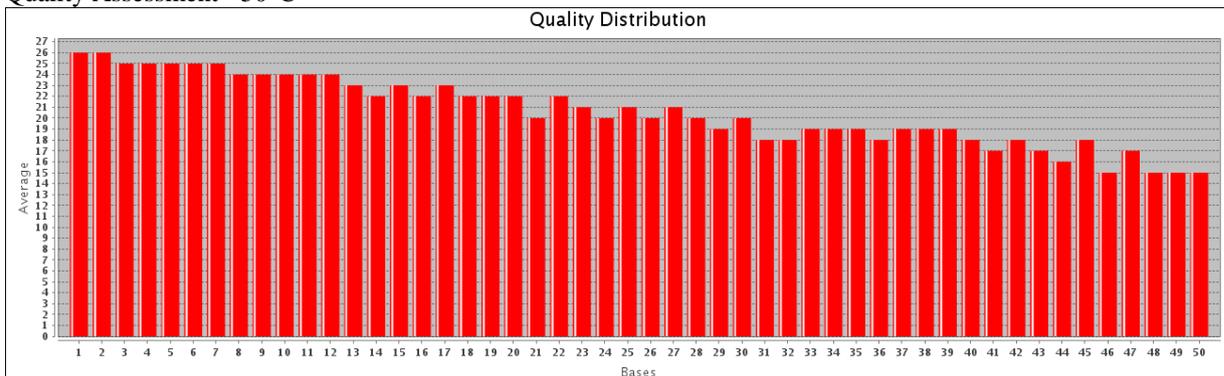


FIGURA 9 - Deposição das *beads* no SOLiD™ em cada condição testada. As *beads* fora da linha tracejada indicam *beads* policlonais e não entraram na análise. pH –condição de estresse ácido, 2M-condição de estresse osmótico, 50°C- condição de estresse térmico e controle

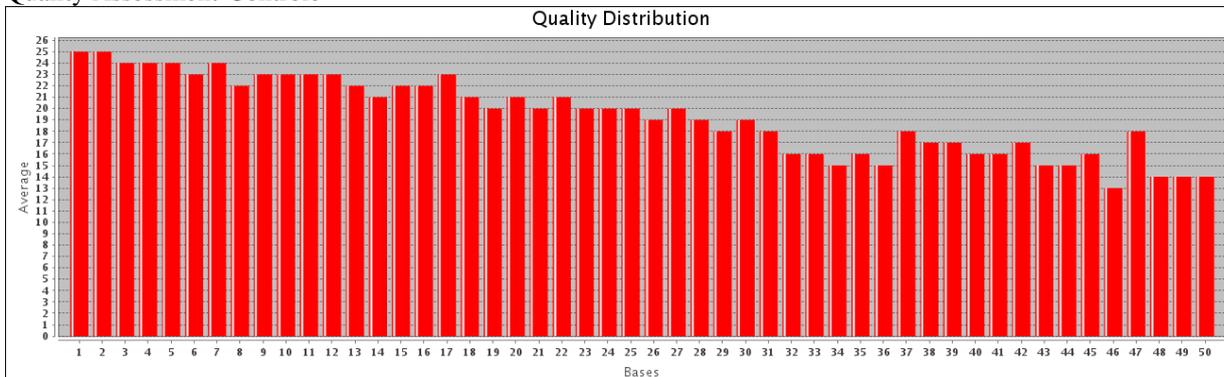
Quality Assessment - 2M



Quality Assessment - 50°C



Quality Assessment-Controle



Quality Assessment- pH

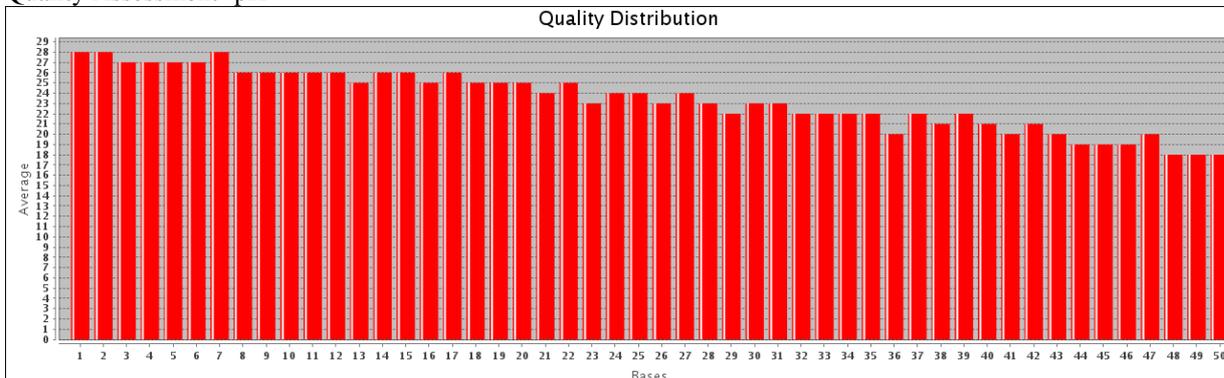


FIGURA 10 - Gráfico representando a distribuição da média de qualidade em cada base do transcrito (50 nt) em cada condição.

Eixo x – cada uma das 50 bases sequenciadas. Eixo y – valor da média de qualidade de cada base. Obtido pelo programa Quality Assessment (Ramos et al., 2011). 2M-estresse osmótico; 50°C-estresse térmico; controle e pH-estresse ácido.

TABELA 1: Quantidade de leituras totais obtidas no sequenciamento.

	2M	50°C	pH	Controle
Dados Brutos	18.783.810	21.622.844	17.393.077	25.235.478
Ribossomais filtrados	9.153.698	9.508.351	9.440.611	8.611.628

Os dados brutos incluem leituras de todos os transcritos. Ribossomais filtrados se refere a quantidade de leituras, referentes aos ribossomais, que não foram consideradas na análise. 2M-estresse osmótico; 50°C-estresse térmico; pH-estresse ácido

O alinhamento foi realizado através do programa Bioscope da Applied Biosystem que permite filtrar os ribossomais que foram transcritos. A tolerância de alinhamento inespecífico foi de no máximo duas bases e as leituras mapeadas em diferentes locais do genoma foram descartadas. Somente leituras únicas foram consideradas. Na tabela 2, temos o total de leituras únicas mapeadas no genoma em cada condição testada. Através dos dados gerados no BioScope inferimos a cobertura, cujo termo é para designar o número de leituras mapeadas em uma posição genômica em particular. Notavelmente, o nível de expressão é específico para cada condição de crescimento. O estresse osmótico foi o mais representado, seguido do estresse ácido, da condição controle e estresse térmico, como pode ser visualizado também na tabela 2.

TABELA 2 - Quantidade de leituras únicas mapeadas no genoma de *C.pseudotuberculosis* 1002 e cobertura dos transcritos no genoma.

	2M	50°C	pH	Controle
Leituras mapeadas unicamente	2.099.781	1.708.863	1.834.898	1.797.169
Cobertura no genoma	52x	43x	46x	45x

2M-estresse osmótico; 50°C-estresse térmico; pH-estresse ácido; condição controle

O programa gera um arquivo de saída com os valores de expressão de cada gene, em uma unidade que já normaliza os dados gerados. No programa Bioscope a normalização é realizada pela medida RPKM (Reads Per Kilobase of coding sequence per Million Mapped), onde a medida de densidade reflete a concentração molar de um transcrito, pela normalização do tamanho do gene e o número total de leitura em cada gene (Mortazavi et al., 2008). O objetivo da normalização é remover os efeitos técnicos que ocorrem nos dados, para assegurar

que estes, tenham o mínimo de impacto nos resultados. No RNAseq, a normalização é padronizada pelo número de leituras no experimento (Robinson e Oshlack, 2010). Segue a fórmula abaixo:

$$\text{RPKM} = \frac{\text{Total de leituras na região transcricionalmente ativa} \times 10^9}{\text{Leituras mapeadas (milhões)} \times \text{Tamanho do gene em pares de bases}}$$

Para análise dos genes diferencialmente expressos utilizamos o pacote DEGseq (Wang, et al., 2009). Este programa utiliza como arquivo de entrada, o arquivo de saída do Bioscope, com valores de RPKM. Para identificar os genes diferencialmente expressos, um corte de significância com o *p-value* < 0,001 foi utilizado, sendo o valor padrão do programa.

O valor da expressão relativa entre estresse e controle é dado através do logaritmo na base 2 (*log2-fold-change*). Os genes classificados como TRUE, ou seja, que foram significativos apresentando um *p-value* < 0,001 foram alvos principais deste trabalho, mas entre estes, fizemos um outro “corte” para analisarmos somente genes que apresentaram a razão de expressão de pelo menos duas vezes (induzido ou reprimido) em relação ao controle. Sendo assim, selecionamos somente genes com o *log2 fold-change* ≥ 1 ou ≤ -1 . Porém, genes envolvidos na virulência ainda que fossem classificados como não significativos (FALSE), descritos no item 1.1.4. foram analisados pois são os genes mais estudados, até o momento, em *C. pseudotuberculosis*. Não vamos utilizar o *log2 fold-change* para visualizarmos os resultados e sim, o *fold-change*, que é a razão entre a expressão dos genes (RPKM) no meio sob estresse, dividido pela expressão de genes (RPKM) no meio controle, pois fica mais fácil visualizar quantas vezes o gene foi mais ou menos expresso em cada situação. O valor de *fold-change* maior ou igual a 2, corresponde ao *log2 fold-change* ≥ 1 e *fold-change* menor ou igual a 0,5, corresponde a *log2 fold-change* ≤ -1 .

5.5. Quantidade de transcritos representados em cada condição testada

Dentre várias perturbações sofridas pela bactéria no processo da infecção, foram simuladas três delas, sendo: estresse osmótico (2M) estresse térmico (50°C) e estresse ácido (pH), além da condição controle. Os transcritos foram analisados no início da fase exponencial ($\text{DO}_{600\text{nm}}=0,2$) e conseguimos representação de todos os transcritos no genoma em todas as condições.

Pela predição do Fgenes (www.softberry.com) e curadoria manual, o genoma da linhagem *C.pseudotuberculosis* 1002 apresentou 2057 regiões codificantes e 48 RNA transportadores somando um total de 2105 regiões no genoma a serem analisadas pelo software. Através do sequenciamento do RNA, realizado pelo SOLiD™, conseguimos representar no início da fase exponencial (Figura 11), 2054 transcritos ativos no controle, equivalendo a 97,58% da região transcrita no genoma, e 51 genes (2,42%) foram considerados não-transcritos, com valor de RPKM igual a zero. Na condição de estresse osmótico, 2064 (98,05%) transcritos, na condição de estresse térmico, 2049 (97,34%) transcritos, e na condição de estresse ácido, 2059 (97,81%) transcritos. Dos 2064 transcritos do estresse osmótico, 991 foram considerados diferencialmente expressos (48,01%) (p -value < 0,001) pelo software de análise DEGseq (Figura 12), em relação ao controle, sendo 741 genes induzidos e 250 genes reprimidos. No estresse térmico, 936 foram considerados diferencialmente expressos (45,68%), sendo 783 induzidos e 153 genes reprimidos e no estresse ácido, 990 genes diferencialmente expressos (48,08%), sendo 745 induzidos e 245 genes reprimidos.

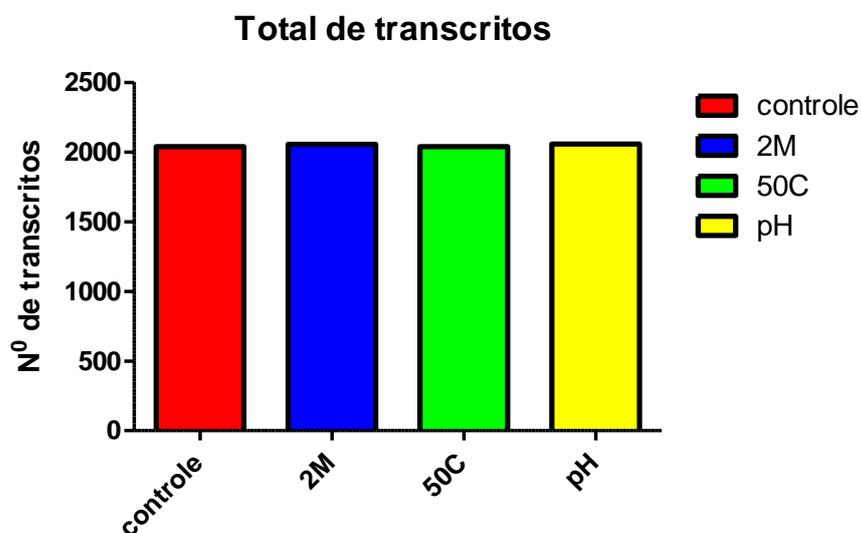


FIGURA 11 - Número total de transcritos ativos que foram confirmados pelo RNAseq através do programa Bioscope.
2M-estresse osmótico; 50°C-estresse térmico; pH-estresse ácido

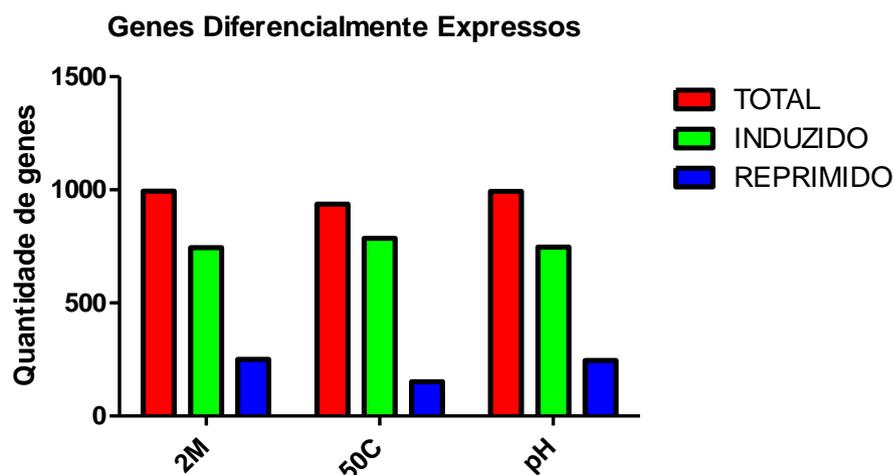


FIGURA 12 – Quantidade de genes considerados diferencialmente expressos em relação ao controle. Vermelho representa todos os genes transcritos em cada condição. Verde – genes considerados induzidos em relação ao controle e Azul - genes considerados reprimidos em relação ao controle. 2M-estresse osmótico; 50°C-estresse térmico; pH-estresse ácido

5.6. Genes envolvidos na virulência

5.6.1. Gene da *pld* e genes *fag*

C. pseudotuberculosis possui poucos genes envolvidos na virulência relatados na literatura. Assim, analisamos a expressão diferencial destes genes em cada situação em que simulamos o ambiente encontrado no hospedeiro, durante o processo de infecção.

Na tabela 3 visualiza-se o valor de expressão do gene codificante da PLD e genes envolvidos na aquisição do ferro (*fag*).

TABELA 3 - Expressão diferencial, em cada condição, do gene *pld*, genes do operon *fagABC* e gene *fagD*.

Condição	Gene	Produto	RPKM estresse	RPKM controle	fold_change
2M	<i>pld</i>	Phospholipase D precursor	1,53	2,36	0,6483*
pH	<i>pld</i>	Phospholipase D precursor	13,44	2,36	5,6949
50°C	<i>pld</i>	Phospholipase D precursor	0	2,36	NA
2M	<i>fagA</i>	integral membrane protein	26,41	44,86	0,5887*
pH	<i>fagA</i>	integral membrane protein	31,85	44,86	0,7099*
50°C	<i>fagA</i>	integral membrane protein	71,16	44,86	1,5862
2M	<i>fagB</i>	Iron-enterobactin transporter	10,48	19,86	0,5276*
pH	<i>fagB</i>	Iron-enterobactin transporter	19,69	19,86	0,9914*
50°C	<i>fagB</i>	Iron-enterobactin transporter	26,25	19,86	1,3217*
2M	<i>fagC</i>	ATP binding cytoplasmic membrane	19,52	22,6	0,8637*

		protein			
pH	<i>fagC</i>	ATP binding cytoplasmic membrane protein	18,06	22,6	0,7991*
50°C	<i>fagC</i>	ATP binding cytoplasmic membrane protein	31,87	22,6	1,4101*
2M	<i>fagD</i>	iron siderophore binding protein	1001,02	1380,39	0,7251*
pH	<i>fagD</i>	iron siderophore binding protein	3707,7	1380,39	2,6859
50°C	<i>fagD</i>	iron siderophore binding protein	1853,52	1380,39	1,3427

RPKM estresse=expressão de genes no meio com alteração. RPKM controle=expressão de genes no meio BHI controle. *fold-change* = valor da razão entre RPKM estresse e RPKM controle, onde valor menor que 1, indica que o gene foi reprimido e valor maior que 1 indica gene induzido. *- genes que não foram significativos, apresentaram *p-value* > 0,001, ou seja apresentou expressão similar ao controle. NA: quando a contagem da leitura para o gene for igual a zero em uma amostra e na outra, for um número menor que 5, o programa não faz cálculo algum para o gene em questão, então aparece NA (Non applicable).

Abaixo seguem as figuras (Figura 13 e Figura 14) obtidas através do programa Artemis do arquivo de saída .bam gerado no programa BioScope, das leituras que cobriram a região no genoma, referente aos genes *pld* e *fag*, entre as condições de estresse ácido, temperatura e controle., respectivamente. Única condição em que o gene *pld* foi considerado diferencialmente expresso foi no meio ácido, ainda que baixo o número de transcritos representando esta região. O gene *fagA*, foi considerado significativo somente no estresse térmico e o *fagD*, significativo no estresse térmico e acidez.

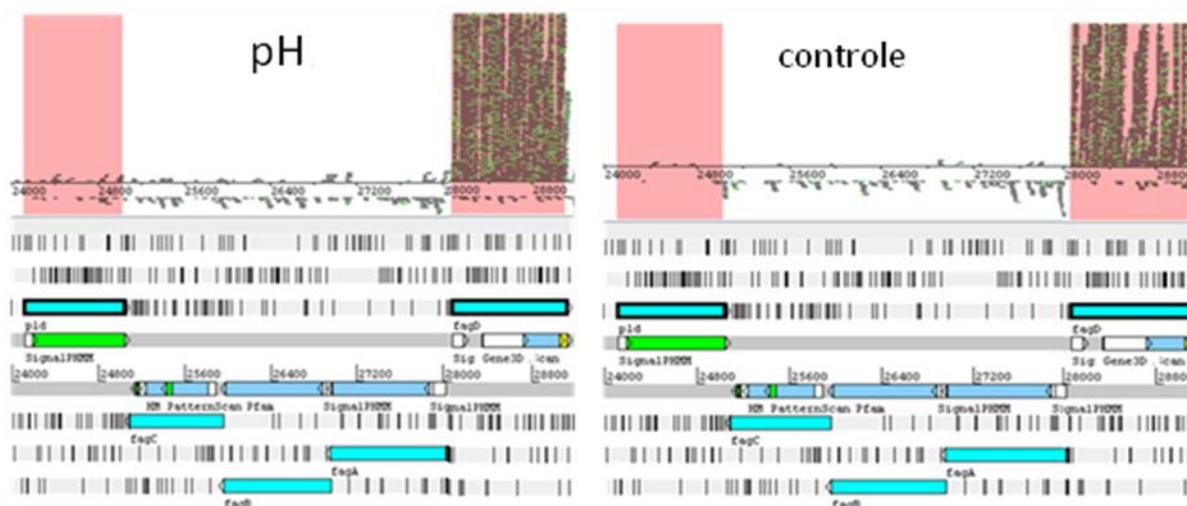


FIGURA 13 - Cobertura da região transcricionalmente ativa do gene *pld* e *fagD*.

Genes *pld* e *fagD*, em destaque cor de rosa, na condição de acidez ao lado esquerdo e condição controle ao lado direito. Observamos os genes considerados diferencialmente expressos na acidez, em relação ao controle. Visualiza-se o número de leituras cobrindo cada região transcricionalmente ativa no genoma de *C.pseudotuberculosis* 1002. Figura obtida no Artemis utilizando o arquivo .bam, gerado no BioScope.

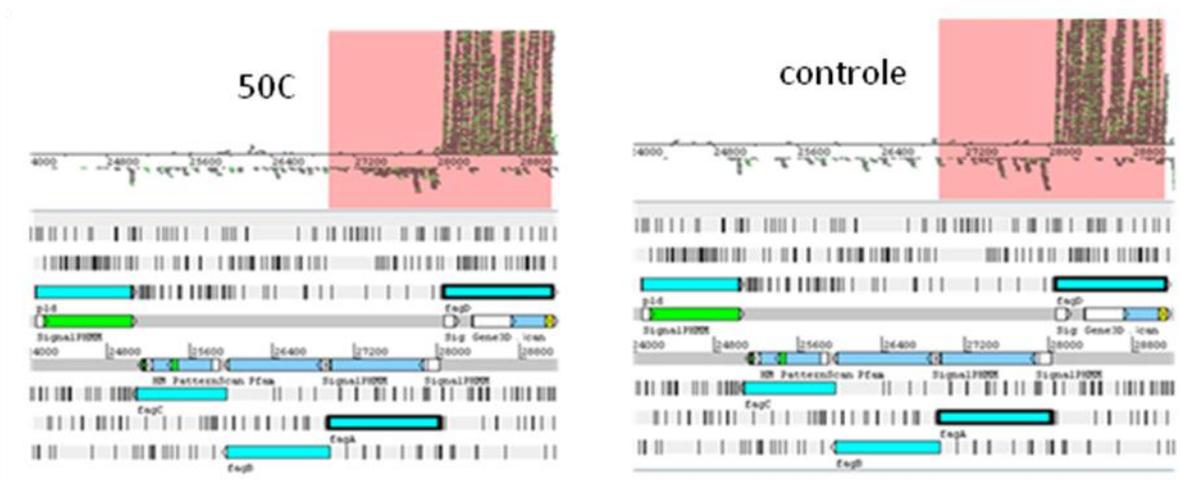


FIGURA 14 - Cobertura da região transcricionalmente ativa do gene *fagA* e *fagD*. Genes *fagA* e *fagD* (destaque em rosa) na condição de estresse térmico à esquerda e condição controle à direita, que foram considerados induzidos no estresse. Observa-se o número de leituras cobrindo cada região transcricionalmente ativa no genoma de *C.pseudotuberculosis* 1002. Figura obtida no Artemis utilizando o arquivo .bam, gerado no BioScope.

Como o gene *fagD* é regulado pelo gene *dtxR*, verificamos os resultados deste gene e os valores podem ser visualizados na tabela 4. O gene foi reprimido nas condições de estresse osmótico e térmico. Na acidez o gene foi considerado induzido.

TABELA 4 - Expressão diferencial do gene *dtxR*.

Condição	Gene	Produto	RPKM estresse	RPKM controle	fold-change	p-value <0,001
2M	<i>dtxR</i>	diphtheria toxin repressor	200,21	348,54	0,5744	0,0052*
50°C	<i>dtxR</i>	diphtheria toxin repressor	347,47	348,54	0,9969	0,0002
pH	<i>dtxR</i>	diphtheria toxin repressor	624,67	348,54	1,7922	4,48 e-46

RPKM estresse=expressão de genes no meio com alteração. RPKM controle=expressão de genes no meio BHI controle, sem alteração. *fold-change* = valor da razão entre RPKM estresse e RPKM controle, onde valor menor que 1, indica que o gene foi reprimido e valor maior que 1 indica gene induzido. *p-value* < 0,001= valor de corte para identificar genes diferencialmente expressos (significativos). * genes que não foram significativos, apresentaram *p-value* > 0,001, ou seja apresentou expressão similar ao controle.

Na figura 15, pode-se observar o regulon do gene *dtxR* e o *fagD*, um dos genes regulados pela DtxR, está em destaque. O gene regula 13 genes, conforme dados do banco CoryneRegNet (<http://coryneregnet.cebitec.uni-bielefeld.de/v6/>).

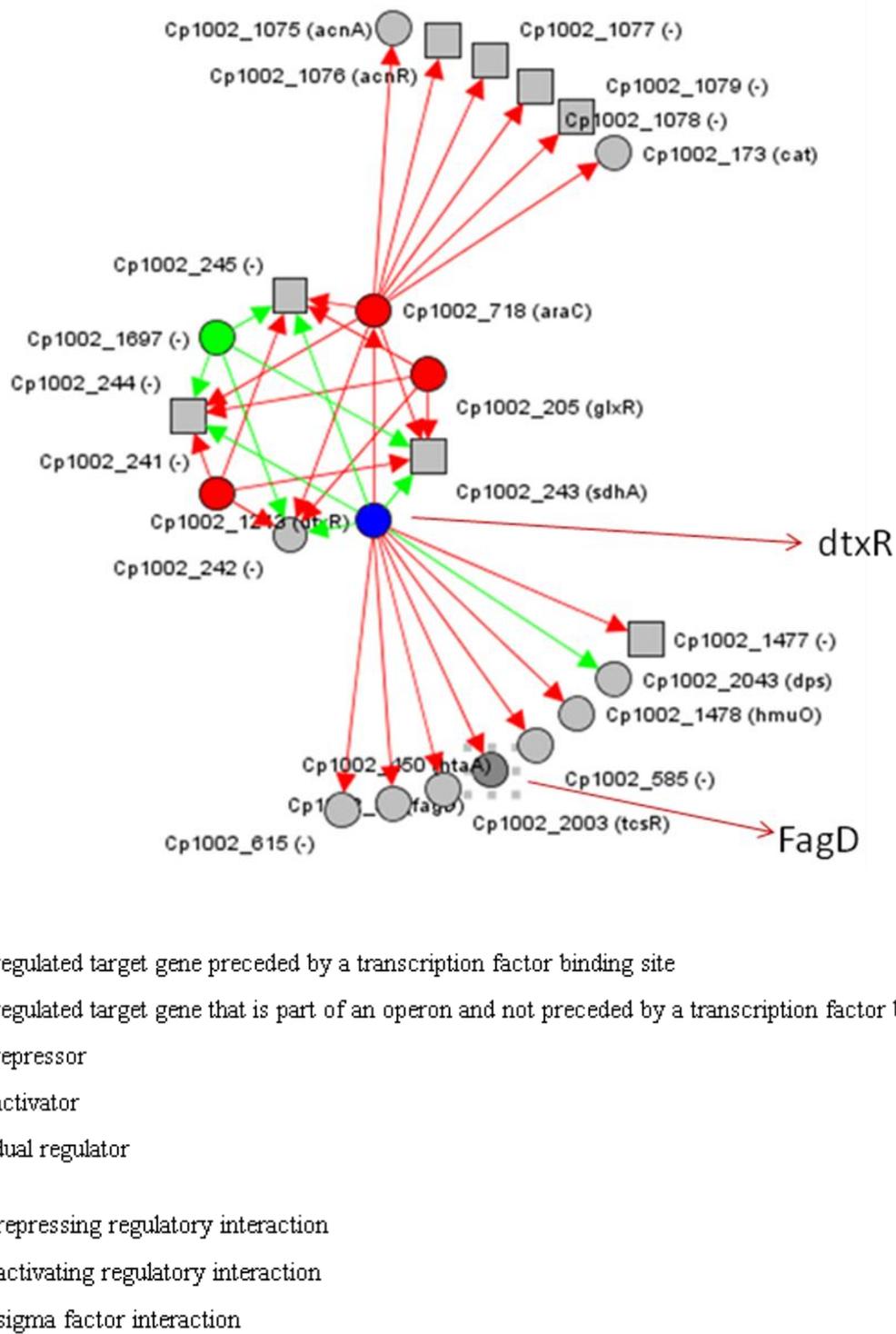


FIGURA 15 - Regulon do gene *dtxR*.

O gene regula 13 genes, entre eles o gene *fagD*. A seta vermelha que sai do círculo azul significa que é uma interação repressora, ou seja o círculo azul que representa o gene *dtxR*, reprime o círculo cinza, que um deles em destaque representa o gene *fagD*. Figura obtida no programa CoryneRegNet (<http://coryneregnet.cebitec.uni-bielefeld.de/v6/>).

5.6.2. Gene *cp40*

Em *C.pseudotuberculosis* 1002, os dados do transcriptoma confirmaram um gene precursor de serina protease- endoS (possivelmente a CP40), mas com uma expressão praticamente nula, muito similar ao controle, em todas as situações testadas. Analisando todo o transcriptoma de *C.pseudotuberculosis* 1002, encontramos quatro supostas proteínas apresentando a característica de uma serina protease, e entre elas, três foram consideradas diferencialmente expressas nas condições testadas (Tabela 5). São elas: *Trypsin family serine protease*, *Putative serine protease*, *Putative trypsin-like serine protease*.

TABELA 5 - Expressão de genes supostamente codificantes de serina protease.

Condição	Gene ID	Produto	RPKM estresse	RPKM controle	Fold-change
2M	Cp1002_0210	Putative Trypsin family serine protease	127,26	71,4	1,7823
pH	Cp1002_0210	Putative Trypsin family serine protease	119,07	71,4	1,6676
50°C	Cp1002_0210	Putative Trypsin family serine protease	81,29	71,4	1,1385*
2M	Cp1002_0662	Putative serine protease	165,9	146,13	1,1352
pH	Cp1002_0662	Putative serine protease	381,33	146,13	2,6095
50°C	Cp1002_0662	Putative serine protease	202,43	146,13	1,3852
2M	<i>sprX</i>	Putative trypsin-like serine protease	144,66	131,51	1,0999
pH	<i>sprX</i>	Putative trypsin-like serine protease	84,51	131,51	0,6426*
50°C	<i>sprX</i>	Putative trypsin-like serine protease	175,6	131,51	1,3352

RPKM estresse=expressão de genes no meio com alteração. RPKM controle=expressão de genes no meio BHI controle, sem alteração. *fold-change* = valor da razão entre RPKM estresse e RPKM controle, onde valor abaixo de 1 indica que o gene foi reprimido e valor acima de 1 indica gene induzido. *P-value* < 0001= valor de corte para identificar genes diferencialmente expressos. * - genes que apresentaram valor de *p-value* > 0,001, ou seja, não foram significativos.

Somente na condição de acidez, o gene Cp1002_0062 codificante da suposta serina protease, apresentou uma indução de quase três vezes (*fold-change* ≥ 2) em relação ao controle.

5.6.3. Componentes da parede celular e modificação lipídica

Em relação à composição dos lipídeos na parede um gene interessante, *ufaA* (Tabela 6), codificante de uma suposta sintase do ácido graxo ciclopropano mostrou variação entre ser classificado como diferencialmente expresso ou não, induzido ou reprimido nas diferentes condições. Somente na condição de acidez o gene não foi considerado diferencialmente expressão e nesta situação mostrou-se reprimido. Nas outras condições foi significativo (p -value < 0,001) e mostrou-se induzido, mas a razão (*fold-change*) entre estresse e controle foi baixa, quase a mesma expressão entre eles.

TABELA 6 - Gene codificante de proteína componente da parede.

Condição	Gene	Produto	RPKM estresse	RPKM controle	<i>Fold-change</i>
2M	<i>ufaA</i>	Putative cyclopropane fatty acid synthase	283,96	216,01	1,3145
pH	<i>ufaA</i>	Putative cyclopropane fatty acid synthase	110,2	216,01	0,5101*
50°C	<i>ufaA</i>	Putative cyclopropane fatty acid synthase	244,73	216,01	1,1329

RPKM estresse=expressão de genes no meio com alteração. RPKM controle=expressão de genes no meio BHI controle, sem alteração. *fold-change* = valor da razão entre RPKM estresse e RPKM controle, onde valor abaixo de 1 indica que o gene foi reprimido e valor acima de 1 indica gene induzido. P-value < 0001=valor de corte para identificar genes diferencialmente expresso. * genes que apresentaram valor de p-value > 0,001, não foram significativos.

Outro gene supostamente envolvido na síntese de lipídeos foi o Cp1002_1953, codificante de uma cutinase, que participa do processo biossintético de ácido micólico e está localizado na fita reversa, como nos mostra a figura 16, *upstream* aos genes (*fadD*, *pks*, e *pccB*) cujo produtos são necessários ao passo final da condensação da biossíntese do ácido micólico (Meniche et al., 2009).



FIGURA 16 - Genes envolvidos na síntese de lipídeos.

Figura obtida no programa Artemis. Genoma de *Corynebacterium pseudotuberculosis* 1002. Da direita para esquerda: Cp1002_1953, *fadD*, *pks*, *pccB*

A tabela 7 mostra os valores de expressão destes genes. Na condição de estresse osmótico e acidez o gene codificante de acyl-CoA sintase (*fadD*) não mostrou uma diferença significativa de expressão em relação ao controle. Em todas as situações estes genes apresentaram *fold-change* baixo, o que equivale, na média, a 1,36 vezes a mais de expressão na condição de estresse em relação ao controle, ou 1,05 vezes a mais na condição controle em relação ao estresse. Os valores de RPKM estão muito próximos, quase não existe uma diferença entre estresse e controle, praticamente a mesma quantidade. Ainda na tabela 7, observa-se que a partir do gene Cp1002_0517, codificante da lipoprotein LpqB, estão os genes que mais expressaram nas condições de estresse em relação ao controle.

TABELA 7 - Genes envolvidos na síntese de lipídeos.

Condição	Gene	Produto	RPKM estresse	RPKM controle	Fold-change
2M	<i>pks</i>	Polyketide synthase	392,57	286,05	1,37
pH	<i>pks</i>	Polyketide synthase	310,9	286,05	1,09
50°C	<i>pks</i>	Polyketide synthase	393,23	286,05	1,37
2M	<i>pccB</i>	Acetyl-CoA carboxylase beta subunit	693,73	536,34	1,29
pH	<i>pccB</i>	Acetyl-CoA carboxylase beta subunit	511,71	536,34	0,95
50°C	<i>pccB</i>	Acetyl-CoA carboxylase beta subunit	529,43	536,34	0,99
2M	<i>fadD</i>	Acyl-CoA synthase	166,13	216,82	0,77*
pH	<i>fadD</i>	Acyl-CoA synthase	194,26	216,82	0,90*

50°C	<i>fadD</i>	Acyl-CoA synthase	267,62	216,82	1,23
2M	Cp1002_1953	Cutinase	626,36	564,37	1,11
pH	Cp1002_1953	Cutinase	658,07	564,37	1,17
50°C	Cp1002_1953	Cutinase	694,99	564,37	1,23
2M	Cp1002_1530	Putative long-chain-fatty-acid-CoA ligase	53,01	99,04	0,54*
pH	Cp1002_1530	Putative long-chain-fatty-acid-CoA ligase	134,37	99,04	1,36
50°C	Cp1002_1530	Putative long-chain-fatty-acid-CoA ligase	155,54	99,04	1,57
2M	Cp1002_0517	lipoprotein LpqB	115,03	48,89	2,3528
pH	Cp1002_0517	lipoprotein LpqB	103,18	48,89	2,1104
50°C	Cp1002_0517	lipoprotein LpqB	85,85	48,89	1,7559
2M	<i>lppL</i>	putative prolipoprotein LppL	388,34	236,28	1,6435
pH	<i>lppL</i>	putative prolipoprotein LppL	184,17	236,28	0,7794*
50°C	<i>lppL</i>	putative prolipoprotein LppL	290,4	236,28	1,229
2M	<i>lgt</i>	Prolipoprotein diacylglyceryl transferase	171,29	78,37	2,1856
pH	<i>lgt</i>	Prolipoprotein diacylglyceryl transferase	91,47	78,37	1,1671*
50°C	<i>lgt</i>	Prolipoprotein diacylglyceryl transferase	121,14	78,37	1,5457
2M	<i>lspA</i>	putative lipoprotein signal peptidase	166,77	202,82	0,8222*
pH	<i>lspA</i>	putative lipoprotein signal peptidase	58,7	202,82	0,2894
50°C	<i>lspA</i>	putative lipoprotein signal peptidase	226,81	202,82	1,1182

RPKM estresse=expressão de genes no meio com alteração. RPKM controle=expressão de genes no meio BHI controle, sem alteração. *fold-change* = valor da razão entre RPKM estresse e RPKM controle, onde valor abaixo de 1 indica que o gene foi reprimido e valor acima de 1 indica gene induzido. *P-value* < 0001=valor de corte para identificar genes diferencialmente expressos. *= genes que apresentaram valor de *p-value* > 0,001, não foram significativos.

5.7. Core estimulon entre as três condições de estresse, com *fold-change* ≥ 2 ou ≤ 0,5

Importante verificar os genes diferencialmente expressos que mostraram uma indução ou repressão de pelo menos duas vezes, em relação ao controle nas três situações de estresse (*core estimulon*), ao mesmo tempo. Assim, para agilizar, foi desenvolvido o programa *DBtransprot* (Sá, dados não publicados) para selecionar os genes de interesse, com base nos dados de transcriptoma previamente inseridos neste banco de dados. Este programa foi utilizado em todos os processos investigados. Os genes que foram induzidos nas três condições de estresse, com o *fold-change* maior ou igual a dois estão descritos na tabela 8. Após o corte estabelecido, obviamente, observou-se que para o mesmo gene analisado se *fold change* foi maior que dois em uma condição, contudo menor que dois em outra, este gene saiu da análise. Porém, como havia muitos genes com valor de RPKM interessante, pelo seu alto valor, criou-

se uma lista com todos os genes comuns, que foram significativos, independente do *fold-change*, nas três condições que estão no anexo 8. Trinta genes foram comuns nas três condições com *fold-change* maior ou igual a dois.

TABELA 8 - Genes induzidos, comuns, nas três condições de estresse, com *fold-change* maior ou igual a dois.

Gene	Produto	<i>Fold-change</i> 2M	<i>Fold-change</i> pH	<i>Fold-change</i> 50°C
Cp1002_1418	Hypothetical protein	33,58	12,85	41,12
Cp1002_0730	TetR family transcriptional regulator	17,8648	6,8803	11,0830
Cp1002_0758	Putative NUDIX family hydrolase	9,2788	5,8134	4,7069
Cp1002_1870	Fimbrial associated sortase	8,5682	2,1690	3,7879
Cp1002_0968	Hypothetical protein	7,7736	3,1734	4,6284
<i>dtd</i>	D-tyrosyl-tRNA(Tyr) deacylase	6,9081	4,4061	5,7842
Cp1002_0025	CRISPR-associated protein	5,9445	2,1775	5,0172
Cp1002_0457	major facilitator superfamily permease	4,9148	2,5621	2,5551
Cp1002_0729	major facilitator superfamily permease	4,6402	2,7056	2,9295
Cp1002_2031	Conserved hypothetical protein	4,3176	4,7596	13,882
Cp1002_0746	Hypothetical protein	4,1618	2,6922	2,8853
Cp1002_0268	Conserved Mitochondrial domain of unknown function	4,1191	2,8433	3,2859
Cp1002_1751	Conserved hypothetical protein	4,1131	2,7604	2,1681
<i>phuC</i>	Iron(III) dicitrate transport permease-like protein yusV	4,0728	4,2327	4,6781
<i>corA</i>	magnesium/cobalt transporter CorA	3,8968	2,6828	2,1770
Cp1002_0757	Conserved hypothetical protein	3,7145	2,3267	2,4797
<i>ylxR</i>	Uncharacterized protein ylxR	3,3564	2,2683	2,1303
<i>ruvC</i>	Crossover junction endodeoxyribonuclease ruvC	3,2668	2,0999	3,6792
<i>hemK</i>	Putative protoporphyrinogen oxidase HemK	3,1215	2,3381	2,7958
<i>rplN</i>	50S ribosomal protein L14	2,9109	2,8606	2,1706
Cp1002_1947	Putative integral membrane protein	2,7826	3,1725	2,3139
Cp1002_1434	Putative oxidoreductase	2,7725	2,7829	5,3759
Cp1002_0893	Putative secreted protein	2,5279	2,2528	2,6139
Cp1002_1134	Hypothetical protein	2,5126	2,5959	2,2542
Cp1002_1547	Putative secreted protein	2,4425	2,0847	2,2664
<i>rplE</i>	50S ribosomal protein L5	2,3546	2,8710	2,0023
<i>rplX</i>	50S ribosomal protein L24	2,3210	2,7409	2,1504
Cp1002_0321	Putative secreted protein	2,1533	2,1684	2,8867
<i>prfA</i>	Peptide chain release factor 1	2,1413	2,1221	2,1245
<i>infA</i>	Translation initiation factor IF-1	2,0392	2,2602	2,3915

Fold-change 2M- razão da expressão entre estresse osmótico e controle. *Fold-change* pH – razão da expressão entre estresse ácido e controle. *Fold-change* 50°C- razão da expressão entre estresse térmico e controle

Entre os genes induzidos, com *fold-change* maior ou igual a 2 e comuns entre as três condições, observamos que 23,3% dos genes são codificantes de proteínas hipotéticas. E o segundo gene mais induzido em todas as três condições, trata-se também de um codificante de

proteína hipotética, Cp1002_1418, cuja expressão ocorreu somente nas condições de estresses, sendo o RPKM deste gene igual a zero na condição controle (anexo 8). Trata-se de um gene codificante de uma proteína com apenas 27 aminoácidos, mas já se mostra importante para estudos futuros.

Entre os genes reprimidos, com valor de *fold-change* pelo menos duas vezes menor ($fold-change \leq 0,5$) que o controle (Tabela 9) observou-se um gene (Cp1002_0008) codificante de uma proteína hipotética, que somente se expressou na condição controle. Em nenhuma condição de estresse o gene foi expresso, e por isso o *fold-change* foi zero. Porém, sua cobertura na condição controle foi baixa, o RPKM foi 18,77 (dado não mostrado). Isto confirma como a tecnologia é sensível em detectar até mesmo transcritos pouco representados no meio.

TABELA 9 - Genes reprimidos comuns entre os três estresses

Gene	Produto	<i>fold-change</i> 2M	<i>fold-change</i> pH	<i>fold-change</i> 50°C
argJ	bifunctional ornithine acetyltransferase/N-acetylglutamate synthase protein	0,4000	0,3906	0,4618
Cp1002_0613	putative secreted protein	0,4176	0,3487	0,4416
Cp1002_1908	Putative oxidoreductase	0,0778	0,0851	0,4357
Cp1002_0573	putative secreted protein	0,2381	0,4038	0,4332
Cp1002_0174	Hypothetical protein	0,2261	0,2361	0,4029
nanK	N-acetylglucosamine kinase	0,2037	0,2898	0,3742
oppA	ABC-type dipeptide/oligopeptide/nickel transport system, secreted component	0,2063	0,2515	0,3525
Cp1002_2063	GntR family regulatory protein	0,2206	0,1880	0,3018
oppB	ABC-type dipeptide/oligopeptide/nickel transport system, permease componen	0,1869	0,3330	0,2999
Cp1002_0008	Hypothetical protein	0	0	0

Fold-change 2M- razão da expressão entre estresse osmótico e controle. *Fold-change* pH – razão da expressão entre estresse ácido e controle. *Fold-change* 50°C- razão da expressão entre estresse térmico e controle

5.8. Estimulon no estresse osmótico, com $fold-change \geq 2$ ou $\leq 0,5$

Para definir o *estimulon*, ou seja, genes diferencialmente expresso em uma condição, fez-se o corte baseado no resultado de DEGseq, que nos informa os genes significativos ou não, em cada situação e também que apresentaram uma razão na expressão de pelo menos duas vezes a mais ou a menos em relação ao controle. Na busca por genes específicos, alguns também foram classificados como comuns na tabela geral (anexo 8). São específicos quando se usa este corte estabelecido, pois apareceu em uma única situação. A tabela 10 mostra a

relação dos genes induzidos na condição de estresse osmótico e a tabela 11, os genes reprimidos. Cento e trinta e seis genes foram considerados induzidos com *fold-change* ≥ 2 , somente no estresse osmótico e setenta e sete genes reprimidos, com *fold-change* $\leq 0,5$.

TABELA 10 - Genes induzidos na condição de estresse osmótico, com *fold-change* ≥ 2

Genes	Produto	RPKM 2M	RPKM controle	<i>fold-change</i>
Cp1002_1667	Hypothetical protein	13,34	0	13,34
Cp1002_0627	Hypothetical protein	39,18	3,02	12,97
Cp1002_1880	Hypothetical protein	10,69	0	10,69
<i>glbO</i>	Group 2 truncated hemoglobin glbO	14,36	1,39	10,33
Cp1002_2050	Hypothetical protein	9,22	0	9,22
Cp1002_0222	Putative secreted protein	14,12	1,63	8,66
Cp1002_0177	Hypothetical protein	34,83	4,03	8,64
Cp1002_0409	Conserved hypothetical protein	24,68	2,86	8,63
Cp1002_0988	Putative secreted protein with LPxTG motif	711,84	99,82	7,13
Cp1002_1313	Putative ATPase with chaperone activity	28,87	5,34	5,41
Cp1002_1616	Putative single-strand binding protein	15,08	2,91	5,18
Cp1002_1612	Putative membrane protein	17,3	3,64	4,75
Cp1002_0015	Hypothetical protein	48,64	10,43	4,66
Cp1002_1613	Conserved hypothetical protein	358,98	84,92	4,23
Cp1002_0824	Hypothetical protein	19,08	4,73	4,03
<i>spaD</i>	Putative fimbrial subunit	644,13	161,85	3,98
<i>vap</i>	Putative virulence-associated protein	1260,54	323,48	3,90
<i>secE</i>	Preprotein translocase subunit SecE	773,52	199,78	3,87
Cp1002_0754	ABC transporter, permease protein	140,66	37,18	3,78
<i>rplU</i>	50S ribosomal protein L21	19454,34	5351,48	3,64
<i>menF</i>	isochorismate synthase	61,34	17,03	3,60
Cp1002_1614	Thioesterase (4HBT) superfamily enzyme	460,29	130,71	3,52
<i>ispF</i>	2-C-methyl-D-erythritol 2,4-cyclo diphosphate synthase	220,38	64,65	3,41
Cp1002_1412	Putative acyltransferase	180,74	53,15	3,40
<i>cumB</i>	Cytosine/adenosine deaminas	173,21	53,72	3,22
<i>smc</i>	Chromosome partition protein	179,93	56,89	3,16
<i>ftsI</i>	Putative penicillin-binding (cell division related) protein	415,9	131,54	3,16
Cp1002_1062	Putative dehydrogenase/oxidoreductase	26,81	8,5	3,15
<i>matE</i>	Putative Multidrug and toxin extrusion protein	117,54	37,34	3,15
<i>rsmE</i>	16S ribosomal RNA methyltransferase RsmE	453,07	144,75	3,13
Cp1002_1417	Putative NPL/P60 family secreted protein	83,58	26,79	3,12
Cp1002_0175	Hypothetical protein	347,98	112,88	3,08
<i>secY</i>	Preprotein translocase subunit SecY	698,29	227,55	3,07
<i>trkA</i>	Trk-type K+ transport system, NAD-binding component	131,67	43,09	3,06
<i>esxT</i>	ESAT-6-like protein	23,34	7,72	3,02
Cp1002_1495	Conserved hypothetical protein	1566,09	519,54	3,01
<i>frgA</i>	Putative siderophore biosynthesis related protein	49,79	16,7	2,98
Cp1002_1874	Putative fimbrial associated sortase-like protein	236,92	80,04	2,96
Cp1002_1176	Hypothetical protein	466,81	158,3	2,95
Cp1002_0958	Hypothetical protein	800,68	273,85	2,92
Cp1002_0290	Putative membrane permease protein, FecCD transport family	20,98	7,24	2,90
<i>udg</i>	UDP-glucose 6-dehydrogenase	113,22	39,12	2,89
Cp1002_0989	Putative surface-anchored membrane protein	177,26	61,25	2,89

continua

Cp1002_1764	Hypothetical protein	196,14	67,91	2,89
tlyA	Putative hemolysin A	570,81	199,35	2,86
ppx2	exopolyphosphatase	858,18	303,89	2,82
yqeL	Putative protein yqeL	757,46	268,71	2,82
Cp1002_0693	putative membrane protein	91,3	32,46	2,81
recB	RecB family nuclease	68,28	24,38	2,80
uppS2	Undecaprenyl pyrophosphate synthase	208,32	74,5	2,80
Cp1002_0088	Conserved hypothetical protein	1063,09	383,51	2,77
Cp1002_0408	phosphoenolpyruvate carboxykinase (ATP)	25,07	9,07	2,76
lysA	diaminopimelate decarboxylase	473,33	172,07	2,75
ectP	Ectoine, glycine betaine and proline transport system	292,66	107,51	2,72
Cp1002_0293	Conserved membrane protein	30,78	11,34	2,71
Cp1002_1048	Conserved hypothetical protein	73,56	27,16	2,71
phoH	Phosphate starvation-inducible protein PhoH	791,73	295,28	2,68
Cp1002_1084	ABC transporter ATP-binding protein	239,97	90,07	2,66
recO	DNA repair protein	119,31	44,81	2,66
Cp1002_1175	Hypothetical protein, UPF0082 protein DIP1378	795,44	300,03	2,65
terC	Integral membrane protein TerC family	200,79	75,78	2,65
proC	Pyrroline-5-carboxylate reductase	212,23	80,35	2,64
nagD	haloacid dehalogenase (HAD) superfamily hydrolase	944,6	357,95	2,64
miaB	(Dimethylallyl)adenosine tRNA methylthiotransferase	418,83	159,89	2,62
Cp1002_1210	Putative methyltransferase	44,34	17,11	2,59
Cp1002_1253	Helix-turn-helix XRE-family like proteins	341,16	132,61	2,57
rpmH	50S ribosomal protein L34	2781,72	1085,02	2,56
rho	transcription termination factor Rho	1231,16	483,41	2,55
Cp1002_0684	Putative ABC transport system	167,6	65,99	2,54
Cp1002_1383	Putative secreted protein	183,24	72,59	2,52
phnA	putative alkyl phosphonate utilization operon protein PhnA	216,24	85,81	2,52
Cp1002_1573	Putative DNA uptake protein, SLBB domain	55,4	22	2,52
hflX	Putative GTP-binding protein hflX	582,17	231,89	2,51
Cp1002_1286	Putative cobalamin synthesis related protein	37,92	15,12	2,51
cdsA	Phosphatidate cytidyltransferase	169,02	67,51	2,50
xerD	site-specific tyrosine recombinase XerD	184,43	74,1	2,49
Cp1002_0605	Putative FK506-binding protein	710,45	285,81	2,49
lysl	L-Lysine transport protein	32,48	13,14	2,47
Cp1002_0924	Membrane protein	48,41	19,65	2,46
Cp1002_1591	7TM-receptor with intracellular phosphohydrolase domain	481,11	195,43	2,46
ppnk	inorganic polyphosphate/ATP-NAD kinase	270,34	111,15	2,43
rnhB	ribonuclease HII	165,17	68,52	2,41
hisF1	Imidazole glycerol phosphate synthase subunit HisF	186,29	77,38	2,41
rimP	Ribosome maturation factor rimP	220,25	92,2	2,39
coaBC	Coenzyme A biosynthesis bifunctional protein coaBC (DNA/pantothenate metabolism flavoprotein)	134,6	56,44	2,38
Cp1002_1457	Putative methyltransferase	49,38	20,72	2,38
purF	Amidophosphoribosyl transferase	911,12	385,92	2,36
Cp1002_0073	Conserved Hypothetical protein	148,38	63,03	2,35
Cp1002_0874	Putative iron ABC transporter membrane protein	190,33	81	2,35
impA	inositol monophosphate phosphatase	94,27	40,33	2,34
yqgF	Putative Holliday junction resolvase	1057,44	452,75	2,34
Cp1002_0815	ATP-dependent helicase	142,44	61,02	2,33
recG	ATP-dependent DNA helicase	30,45	13,13	2,32
pgsA1	Putative Phosphatidylglycerophosphate synthase	231,12	99,9	2,31
holA	DNA polymerase III delta subunit	39,06	16,96	2,30

<i>arc</i>	putative ATPase Arc	65,32	28,41	2,30
Cp1002_0543	Putative membrane transport protein	186,32	81,16	2,30
<i>rpmE</i>	50S ribosomal protein L31 type B	2474,01	1080,64	2,29
Cp1002_1148	Uncharacterized AAA domain-containing protein Rv2559c/MT2636	93,12	40,69	2,29
<i>glk</i>	glucose kinase	191,21	83,89	2,28
Cp1002_1398	Putative membrane protein	784,55	345,01	2,27
<i>rpsP</i>	30S ribosomal protein S16	5139,29	2286,48	2,25
<i>trkH</i>	TrkH family potassium uptake protein	166,22	74,22	2,24
<i>sigA</i>	RNA polymerase sigma factor rpoD (Sigma-A)	582,09	259,94	2,24
mfsD1	Major facilitator superfamily domain-containing protein 1	57,2	25,61	2,23
Cp1002_1183	Putative lipid A biosynthesis acyltransferase	194,62	87,18	2,23
<i>ribF</i>	Putative Riboflavin biosynthesis protein ribF	203,78	91,55	2,23
<i>phoB</i>	Two-component regulatory protein	105,15	47,31	2,22
<i>ada</i>	methylated-DNA--protein-cysteine methyltransferase	89,92	40,66	2,21
<i>trmA</i>	Putative RNA methyltransferase, TrmA family	191,27	86,61	2,21
<i>lgt</i>	Prolipoprotein diacylglyceryltransferase	171,29	78,37	2,19
Cp1002_1727	Hemolysin III-like protein	36,75	16,86	2,18
<i>nusA</i>	Transcription elongation protein	362,85	167,3	2,17
<i>cinA</i>	Putative competence-damage inducible protein	321,63	150,87	2,13
<i>uvrC</i>	UvrABC system protein C	97,77	46,31	2,11
<i>gpmB</i>	Phosphoglycerate mutase family protein	384,21	183,66	2,09
Cp1002_0767	Putative GTP-binding elongation factor	454,65	217,93	2,09
Cp1002_0962	Putative thiamin pyrophosphokinase, catalytic domain protein	694,99	333,3	2,09
Cp1002_0054	Hypothetical protein	254,2	122,88	2,07
Cp1002_1415	Putative glycosyltransferase	113,48	55,35	2,05
<i>lepB1</i>	Signal peptidase I	188,41	92,07	2,05
Cp1002_2010	Putative membrane protein	137,35	67,26	2,04
<i>relA</i>	GTP pyrophosphokinase	562,94	275,78	2,04
<i>dhh</i>	Phosphoesterase DHH family protein	495,24	242,87	2,04
<i>recR</i>	Recombination protein RecR	103,05	50,55	2,04
<i>dnaJ2</i>	Chaperone protein dnaJ 1	713,29	350,01	2,04
Cp1002_1626	ABC transporter ATP-binding protein	95,66	47,13	2,03
<i>pheS</i>	phenylalanyl-tRNA synthetase subunit alpha	387,95	191,49	2,03
<i>nrfC</i>	Cytochrome c nitrate reductase, small subunit	423,71	209,47	2,02
Cp1002_1789	Conserved hypothetical protein	93,61	46,28	2,02
Cp1002_0670	Putative acetyltransferase	118,2	58,44	2,02
Cp1002_1180	Adenosylmethionine--8-amino-7-oxononanoate aminotransferase	35,18	17,46	2,01
<i>infB</i>	Translation initiation factor IF-2	1365,16	678,71	2,01
Cp1002_0120	Di- and tricarboxylate transporter	46,57	23,16	2,01
Cp1002_0431	Two component LuxR family transcriptional	69,34	34,53	2,01
<i>scpA</i>	segregation and condensation protein A	145,24	72,45	2,00

RPKM 2M- valor da expressão do gene no estresse osmótico e RPKM controle-- valor da expressão do gene no controle. *Fold-change*– razão da expressão entre estresse osmótico e controle.

Para facilitar a visualização da razão de expressão entre os genes reprimidos, o cálculo do *fold-change* é mostrado de duas maneiras: divisão entre estresse/controle e divisão entre controle/estresse, pois assim, fica mais fácil visualizar quantas vezes o controle está acima do estresse.

TABELA 11 - Genes reprimidos na condição de estresse osmótico, com *fold-change* ≤ 0,5.

Gene	Produto	RPKM 2M	RPKM controle	<i>fold- change</i>	<i>fold-change</i> controle/estresse
Cp1002_0438	Putative secreted protein	0	13,89	0	13,89
<i>cdd</i>	Cytidine deaminase	5,99	63,57	0,094	10,613
<i>deoA</i>	thymidine phosphorylase	2,81	24,82	0,113	8,833
<i>araC</i>	Putative AraC-family transcriptional regulator	47,52	385,15	0,123	8,105
<i>mntA</i>	Manganese ABC transporter substrate-binding protein	6,88	51,2	0,134	7,442
Cp1002_1580	Putative membrane protein	13,67	97,43	0,140	7,127
<i>deoD</i>	Purine nucleoside phosphorylase	15,08	78,21	0,193	5,186
<i>ald</i>	Alanine dehydrogenase	35,58	178,52	0,199	5,017
Cp1002_0519	ribosome-associated protein Y	926,82	4582,65	0,202	4,944
<i>yihR</i>	Putative aldose 1-epimerase	77,83	369,06	0,211	4,742
Cp1002_0454	Putative membrane protein	20,17	90,43	0,223	4,483
Cp1002_1342	Putative UTP--hexose-1-phosphate uridylyltransferase	9,44	39,29	0,240	4,162
Cp1002_1323	Conserved hypothetical protein	12,12	49,58	0,244	4,091
Cp1002_1975	GntR family transcriptional regulator	37,76	149,4	0,253	3,957
Cp1002_1876	Putative ABC transport system,ATP-binding subunit	12,82	50,59	0,253	3,946
Cp1002_0505	Conserved hypothetical protein	11,34	44,17	0,257	3,895
Cp1002_0079	Conserved hypothetical protein	35,09	130,01	0,270	3,705
Cp1002_1603	Hypothetical protein	26,12	95,05	0,275	3,639
Cp1002_0343	Hypothetical protein	24,98	89,42	0,279	3,580
Cp1002_1962	Conserved hypothetical protein	44,03	154,97	0,284	3,520
Cp1002_0609	Putative oxidoreductase	139,79	489,39	0,286	3,501
<i>arcB</i>	Ornithine cyclodeaminase	15,33	53,24	0,288	3,473
<i>yutK</i>	Nucleoside transporter	36,76	125,46	0,293	3,413
<i>dhaK</i>	Putative dihydroxyacetone kinase subunit	35,66	121,16	0,294	3,398
Cp1002_0342	Conserved hypothetical protein	30,53	94,27	0,324	3,088
<i>oppF</i>	ABC-type transporter, ATPase component	29,31	90,08	0,325	3,073
<i>cat1</i>	Succinyl-CoA Coenzyme A transferase	97,71	297,5	0,328	3,045
<i>mmdA</i>	methylmalonyl-CoA decarboxylase, alpha-subunit	466,53	1341,59	0,348	2,876
<i>sseA</i>	Thiosulfate sulfurtransferase	78,63	226,06	0,348	2,875
<i>fruK</i>	Putative 1-phosphofructokinaseB	88,18	250,83	0,352	2,845
<i>thiX</i>	Thiamine biosynthesis protein X	134,91	380,71	0,354	2,822
<i>furB</i>	Putative ferric uptake regulatory protein	197,55	544,4	0,363	2,756
Cp1002_1852	Putative secreted protein	78,36	215,73	0,363	2,753
Cp1002_1467	Putative cell-surface hemin receptor	30,37	82,59	0,368	2,719
Cp1002_0315	Putative surface-anchored membrane protein	83,78	227,69	0,368	2,718
Cp1002_0094	Conserved hypothetical protein	44,78	120,77	0,371	2,697
Cp1002_0585	Putative Metal binding protein TroA_a	999,48	2640,86	0,378	2,642
<i>nagA1</i>	N-acetylglucosamine-6-phosphate deacetylase	32,24	85,06	0,379	2,638
Cp1002_0623	ABC transporter solute-binding protein	77,59	203,13	0,382	2,618

<i>hmuO</i>	Heme oxygenase	44,88	115,87	0,387	2,582
Cp1002_2058	Substrate-binding protein	121,26	312,14	0,388	2,574
Cp1002_2073	Conserved hypothetical protein	66,22	169,54	0,391	2,560
<i>sodA</i>	Manganese superoxide dismutase	315,77	802,61	0,393	2,542
Cp1002_0303	Putative lactate dehydrogenase	55,17	139,63	0,395	2,531
<i>uspA</i>	universal stress protein	782,57	1966,86	0,398	2,513
<i>maf</i>	Maf-like protein	53,73	133,08	0,404	2,477
<i>gcvP</i>	Glycine cleavage system P protein	53,17	131,33	0,405	2,470
Cp1002_2089	Putative iron-sulphur protein	63,53	155,72	0,408	2,451
Cp1002_0610	enoyl-CoA hydratase	101,44	247,58	0,410	2,441
Cp1002_0086	Serine 3-dehydrogenase	84,39	202,4	0,417	2,398
Cp1002_0345	Alkaline shock protein	113,52	272,2	0,417	2,398
Cp1002_1540	Putative secreted protein	146,46	351,05	0,417	2,397
Cp1002_2069	Putative secreted protein	189,18	451,07	0,419	2,384
<i>menB</i>	Naphthoate synthase	49	116,3	0,421	2,373
Cp1002_0571	Hypothetical protein	124,66	294,88	0,423	2,365
Cp1002_0095	Metalloendopeptidase	55	127,96	0,430	2,327
Cp1002_0833	Hypothetical protein	67,77	156,94	0,432	2,316
Cp1002_0602	Putative DNA-binding protein	205,2	471,21	0,435	2,296
Cp1002_0640	Putative succinate-semialdehyde dehydrogenase	224,65	513,3	0,438	2,285
<i>nagB</i>	Glucosamine-6-phosphate isomerase	66,3	151,45	0,438	2,284
Cp1002_0937	Putative zinc (Zn ²⁺)-dependent hydrolase	64,49	147,11	0,438	2,281
<i>deoR1</i>	Deoxyribose operon repressor	383,6	874,83	0,438	2,281
<i>cobG</i>	Putative precorrin-3B synthase	300,24	678,18	0,443	2,259
Cp1002_0572	Putative decarboxylase	388,55	862,34	0,451	2,219
<i>gltA</i>	type II citrate synthase	67,48	149,12	0,453	2,210
Cp1002_0344	Conserved hypothetical protein	76,28	165,41	0,461	2,168
<i>dctA</i>	Aerobic C4-dicarboxylate transport protein	131,36	282,19	0,466	2,148
Cp1002_0048	Conserved hypothetical protein	512,16	1098,15	0,466	2,144
<i>fumC</i>	fumarate hydratase class II	351,61	749,13	0,469	2,131
<i>copC</i>	Putative Copper resistance protein CopC	86,97	185,27	0,469	2,130
Cp1002_0512	conserved hypothetical protein	136,79	289,12	0,473	2,114
<i>malk</i>	Glycerol-3-phosphate-transporting ATPase	139,76	293,99	0,475	2,104
<i>mgo</i>	malate:quinone oxidoreductase	664,79	1385,32	0,480	2,084
Cp1002_0071	Transcriptional regulator protein	201,43	419,51	0,480	2,083
<i>pgi</i>	glucose-6-phosphate isomerase	92,37	189,41	0,488	2,051
<i>glyA</i>	Serine hydroxymethyltransferase	421,68	852,89	0,494	2,023
Cp1002_1370	Conserved hypothetical protein	289,32	581,62	0,497	2,010

RPKM 2M- valor da expressão do gene no estresse osmótico e RPKM controle- - valor da expressão do gene no controle. *Fold-change*– razão da expressão entre estresse osmótico e controle. *Fold-change* controle/estresse- razão da expressão entre a condição controle e estresse osmótico, somente para facilitar a visualização da repressão do gene no estresse.

5.8.1. Processo Biológico do estímulo dos genes induzido no meio osmótico

Através do programa Blast2GO, identificamos os processos mais abundantes na célula entre os genes induzidos durante a permanência no estresse osmótico. No intuito de facilitar a identificação de quais genes estariam compondo os processos de interesse, foi desenvolvido o programa CoreStImulon (CSI) (figura 17) que permitiu então, identificar os genes presentes nas categorias do Blast2GO, cuja lista se encontra no (anexo 9). O CSI é baseado nos resultados de Blast2GO, contudo as categorias são listadas em um nível muito mais específico, sem muitos agrupamentos de genes relacionados, como ocorre no Blast2GO nível 3, por exemplo. Pelo CSI, observa-se que no estresse osmótico, alguns processos menos representados parecem de maior relevância ao processo infeccioso, como o processo de adesão, composto por 2 genes (Cp1002_0988= secreted protein with I motif e Cp1002_0989 =surface-anchored membrane protein). Foram selecionados os 20 primeiros processos para serem visualizados nos gráficos abaixo. Entre estes, o processo metabólico e DNA de reparo foram os mais representados.

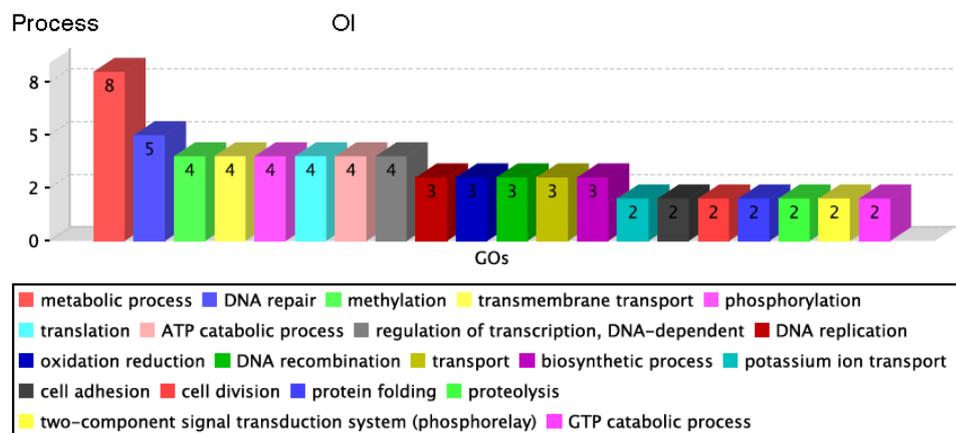


FIGURA 17 - Frequências de genes induzidos pertencentes ao estímulo osmótico. Figura obtida pelo programa CSI.

5.8.2. Processo Biológico do estímulo dos genes reprimidos no meio osmótico

Na figura 18, observa-se que entre os processos mais representados estão aqueles envolvidos com o crescimento da bactéria. Pelo CSI, visualizamos os processos metabólicos e ciclo do ácido tricarbóxico reduzido, justificando o crescimento desacelerado da bactéria

durante o tempo que permaneceu no estresse neste ambiente. Os gene que compõe cada processo podem ser visualizados no anexo 10.

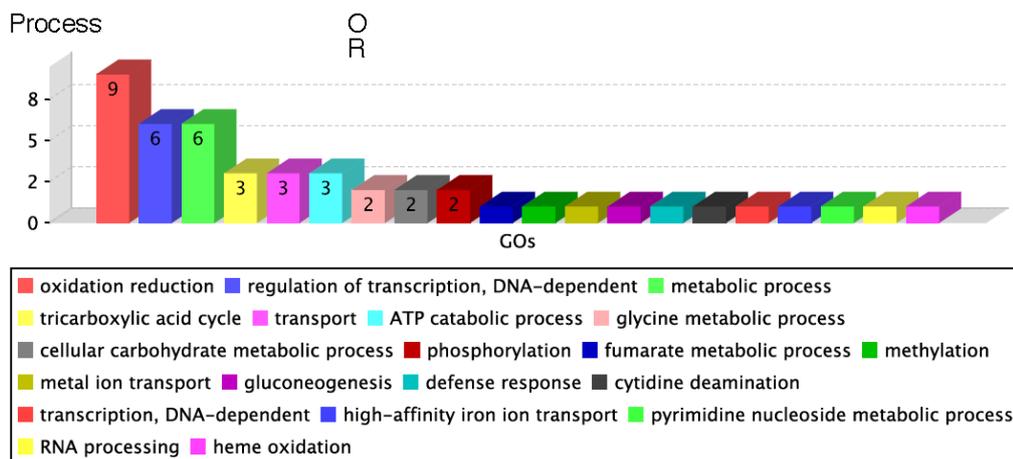


FIGURA 18 - Frequências de genes reprimidos pertencentes ao estimulon osmótico. Figura obtida pelo programa CSI.

5.9. Estimulon do estresse por acidez, com $fold-change \geq 2$ ou $\leq 0,5$

Na condição de estresse ácido, cento e sete genes foram destacados com indução de um $fold-change \geq 2$ e setenta e dois genes reprimidos, com $fold-change \leq 0,5$, como podem ser observados nas tabelas 12 e 13, respectivamente.

TABELA 12 - Genes induzidos na condição de estresse ácido, com $fold-change \geq 2$.

Gene	Produto	RPKM pH	RPKM controle	$fold-change$
<i>msrB</i>	Peptide methionine sulfoxide reductase msrB	671,2	33,11	20,27
Cp1002_1082	Hypothetical protein	48,45	3,49	13,88
Cp1002_0084	Alkylated DNA repair protein	11,61	0	11,61
<i>nrtd1</i>	Ribonucleotide reductase stimulatory protein	11,41	0	11,41
Cp1002_2051	Hypothetical protein	10,91	0	10,91
<i>padR</i>	PadR family transcriptional regulator	304,92	30,24	10,08
Cp1002_0218	Putative type II secretion system protein	9,44	0	9,44
Cp1002_2017	Putative GntR-family transcriptional regulator	238,97	29,03	8,23
<i>dps</i>	DNA protection during starvation protein	12269,57	1493,78	8,21
Cp1002_1558	Hypothetical protein	32,72	4,12	7,94
Cp1002_0023	Conserved hypothetical protein	24,32	3,27	7,44
<i>lysR</i>	Transcriptional activator protein lysR	1437,89	199,38	7,21
<i>acnA</i>	Aconitate hydratase	853,33	126,56	6,74
Cp1002_1624	YcaO-like family protein	63,64	9,59	6,64
<i>pld</i>	Phospholipase D precursor	13,44	2,36	5,69

continua

Cp1002_0215	Hypothetical protein	24,99	5,04	4,96
Cp1002_1984	Cell-surface hemin receptor	67,58	14,13	4,78
Cp1002_0041	Conserved hypothetical protein	131,15	28,71	4,57
Cp1002_2016	ABC transporter ATP-binding protein	126,52	28,49	4,44
<i>pdxR</i>	HTH-type pyridoxine biosynthesis transcriptional regulator PdxR	15,48	3,51	4,41
<i>mgtE</i>	Mg ²⁺ transporter mgtE	29,28	6,69	4,38
<i>cat</i>	Catalase	1332,97	305,98	4,36
Cp1002_1604	Putative membrane protein	219,55	50,45	4,35
<i>deoD</i>	Purine nucleoside phosphorylase	339,57	78,21	4,34
Cp1002_1964	Putative secreted protein with LPxTG motif	24,46	5,64	4,34
Cp1002_0562	putative secreted protein	22,39	5,65	3,96
<i>marR</i>	Putative MarR-family regulatory protein	260,22	66,58	3,91
Cp1002_1638	Putative oxidoreductase	115,64	29,97	3,86
Cp1002_1463	Hypothetical protein	55,15	14,63	3,77
<i>nrdG</i>	Anaerobic ribonucleoside-triphosphate reductase activating protein	44,01	11,89	3,70
Cp1002_0161	MFS family major facilitator transporter, tetracycline:cation symporter tetracycline	46,55	12,62	3,69
<i>sigH</i>	RNA polymerase sigma factor RpoE	1175,3	320	3,67
Cp1002_1685	Hypothetical protein	17,26	4,97	3,47
Cp1002_1477	Conserved hypothetical protein	179,95	52,15	3,45
Cp1002_1607	DSBA-like thioredoxin domain-containing protein	2041,18	606,64	3,36
Cp1002_1402	Putative transmembrane protein	498,95	148,47	3,36
Cp1002_1935	Putative secreted protein	351,33	106,58	3,30
Cp1002_2053	ABC transporter inner membrane protein	63,06	19,65	3,21
Cp1002_0592	Putative ATP-dependent DNA helicase	18,98	5,94	3,20
Cp1002_1682	ArsR family transcriptional regulator	64,94	20,47	3,17
Cp1002_2002	LSR2-like protein	1015,8	320,53	3,17
<i>cgt54</i>	two component sensor kinase	1687,96	533,41	3,16
<i>mntB</i>	Manganese ABC transporter ATP-binding protein	28,37	9,04	3,14
<i>sigB</i>	RNA polymerase sigma factor sigB	1187,13	384,38	3,09
Cp1002_0526	Putative GTPase EngC	105,63	34,26	3,08
Cp1002_0076	Putative membrane protein	173,9	56,42	3,08
Cp1002_1623	Putative membrane protein	23,66	7,77	3,05
<i>menC</i>	O-succinylbenzoate synthase	280,33	93,36	3,00
Cp1002_2001	Conserved hypothetical protein	1300,59	433,46	3,00
Cp1002_0898	Putative integral membrane protein	493,2	164,66	3,00
Cp1002_1471	Conserved hypothetical protein	1529,58	510,7	3,00
<i>yxIE</i>	Integral membrane protein YxIE	35,59	11,96	2,98
<i>pgo</i>	Pyruvate dehydrogenase	71,17	24,76	2,87
Cp1002_1008	Putative SAM-dependent methyltransferase	50,5	17,59	2,87
<i>leuD</i>	3-isopropylmalate dehydratase small subunit	21,01	7,37	2,85
<i>mraZ</i>	Cell division protein MraZ	3866,44	1363,52	2,84
Cp1002_0289	Putative ABC-type transporter	29,54	10,48	2,82
<i>tatB</i>	Sec-in dependent twin-arginine translocase	579,98	207,22	2,80
Cp1002_1936	Arabinofuranosyltransferase	42,84	15,34	2,79
Cp1002_2052	BC transport system ATP-binding protein	26,05	9,55	2,73
Cp1002_0077	Conserved hypothetical protein	97,39	35,77	2,72
Cp1002_1659	Rhomboid family protein	398,13	147,57	2,70
<i>fagD</i>	iron siderophore binding protein-FagD	3707,7	1380,39	2,69
Cp1002_1910	Conserved hypothetical protein	264,59	100,26	2,64
Cp1002_1560	Putative transport system membrane protein	21,33	8,11	2,63
Cp1002_0795	Conserved hypothetical protein	267,23	101,62	2,63
Cp1002_1403	Conserved hypothetical protein	232,28	88,8	2,62
Cp1002_0662	Putative serine protease	381,33	146,13	2,61
Cp1002_0858	Conserved hypothetical protein	429,39	165,3	2,60
Cp1002_1469	Inosine 5-monophosphate dehydrogenase	358,77	138,19	2,60
<i>pulA</i>	Putative pullulanase type I	43,48	16,84	2,58
<i>sigE</i>	RNA polymerase sigma factor SigE	729,05	283,22	2,57
<i>arcB3</i>	Putative ornithine cyclodeaminase	22,43	8,83	2,54

<i>ptpA</i>	Low molecular weight protein-tyrosine-phosphatase	137,23	54,83	2,50
Cp1002_1474	Hypothetical protein	298,64	119,69	2,50
<i>yihR</i>	Putative aldose 1-epimerase	913,74	369,06	2,48
<i>gntK</i>	Gluconokinase	57,06	23,24	2,46
Cp1002_0498	Conserved hypothetical protein	62,9	25,69	2,45
<i>dinB</i>	DNA polymerase IV	82,11	33,66	2,44
<i>ureA</i>	Urease gamma subunit	39,2	16,17	2,42
<i>ruvA</i>	Holliday junction ATP-dependent DNA helicase ruvA	377,54	157,45	2,40
Cp1002_1083	Hypothetical protein	24,23	10,14	2,39
<i>piuB</i>	PepSY-associated TM helix	105,12	44,17	2,38
Cp1002_0677	Putative membrane protein	500,02	212,12	2,36
<i>srmB</i>	Putative Superfamily II DNA and RNA helicases	82,13	34,96	2,35
<i>gltA</i>	type II citrate synthase	349,49	149,12	2,34
<i>argK</i>	Putative Arginine/ornithine transport system ATPase	268,45	115,52	2,32
<i>araC</i>	Putative AraC-family transcriptional regulator	892,41	385,15	2,32
<i>cstA</i>	Carbon starvation protein CstA	148,31	64,06	2,32
Cp1002_0797	putative secreted protein	269,52	116,6	2,31
<i>hmuO</i>	Heme oxygenase	256,83	115,87	2,22
<i>rpIW</i>	50S ribosomal protein L23	2217,63	1040,76	2,13
<i>cseE</i>	anti-sigma factor	1067,21	504,07	2,12
Cp1002_0380	Putative ABC transport system membrane protein	35,55	16,83	2,11
<i>ino1</i>	Myo-inositol-1-phosphate synthase	50,7	24,06	2,11
Cp1002_0250	Putative oxidoreductase	390,31	185,85	2,10
<i>ctpA</i>	Putative cation-transporting P-type ATPase	47,09	22,71	2,07
Cp1002_0199	atB/YqeY domain protein	410,04	198,62	2,06
<i>mnmA</i>	tRNA (5-methyl aminomethyl-2-thiouridylate)- methyltransferase	214,69	104,11	2,06
Cp1002_2018	Conserved hypothetical protein	618,31	299,95	2,06
Cp1002_0571	Hypothetical protein	607,33	294,88	2,06
Cp1002_0019	Conserved hypothetical protein	1427,56	694,27	2,06
<i>mntA</i>	Manganese ABC transporter substrate-binding protein	103,8	51,2	2,03
<i>pnuC</i>	Putative nicotinamide mononucleotide transporter PnuC	35,69	17,61	2,03
Cp1002_1774	Conserved hypothetical protein	52,49	26,04	2,02
Cp1002_1565	Hypothetical protein	60,89	30,24	2,01

RPKM pH- valor de expressão do gene no meio ácido. RPKM controle-valor de expressão do gene no meio controle. *Fold-change* – razão do valor de RPKM pH/RPKM controle.

TABELA 13 - Genes reprimidos na condição de acidez, com *fold-change* $\leq 0,5$

Gene	Product	RPKM pH	RPKM controle	<i>fold- change</i>	<i>fold-change controle/estresse</i>
Cp1002_1888	Hypothetical protein	0	13,28	0	13,28
Cp1002_1302	Putative membrane protein	67,16	410,22	0,164	6,11
Cp1002_0922	Putative ABC transport system membrane protein	10,15	56,32	0,180	5,55
Cp1002_1408	Putative transcription regulator	38,13	210,69	0,181	5,53
Cp1002_1314	Conserved hypothetical protein	81,15	400,29	0,203	4,93
<i>odhl</i>	Putative inhibitor of ODH activity	198,7	807,78	0,246	4,07
Cp1002_1241	Hypothetical protein	37,52	150,58	0,249	4,01
Cp1002_1602	Hypothetical protein	13,43	51,46	0,261	3,83
Cp1002_0432	Conserved hypothetical protein	34,97	131,58	0,266	3,76

nusG	Transcription anti-termination protein NusG	162,93	604,48	0,270	3,71
lspA	Putative lipoprotein signal peptidase	58,7	202,82	0,289	3,46
ppiA	Peptidyl-prolyl cis-trans isomerase A	298,58	1027,97	0,290	3,44
argC	N-acetyl-gamma-glutamyl-phosphate reductase	31,76	105,06	0,302	3,31
Cp1002_1231	Hypothetical protein	286,08	944,56	0,303	3,30
rpfB	resuscitation-promoting factor RpfB	68,42	219,27	0,312	3,20
Cp1002_0886	Putative low molecular weight protein antigen 6	45,99	147,19	0,312	3,20
Cp1002_1126	Hypothetical protein	470,08	1496,17	0,314	3,18
ppiB	Putative peptidyl-prolyl cis-trans isomerase B	52,41	166,19	0,315	3,17
Cp1002_0415	Conserved hypothetical protein	102,99	324,13	0,318	3,15
gcvH	glycine cleavage system H protein	48,82	153,33	0,318	3,14
Cp1002_1248	Conserved hypothetical protein	22,81	71,56	0,319	3,14
pyrH	uridylyate kinase	328,19	1007,73	0,326	3,07
cinA	Putative competence-damage inducible protein	50,22	150,87	0,333	3,00
Cp1002_0731	Conserved hypothetical protein	217,55	652,74	0,333	3,00
Cp1002_1684	Hypothetical protein	37,03	107,54	0,344	2,90
argH	argininosuccinase lyase	225,31	653,73	0,345	2,90
Cp1002_1867	Putative surface-anchored protein	103,86	296,22	0,351	2,85
Cp1002_0997	Putative MerR-family transcription regulator	266,7	756,95	0,352	2,84
Cp1002_0954	Conserved hypothetical protein	230,88	650,54	0,355	2,82
fepC	ABC transporter ATP-binding protein	195,27	537,96	0,363	2,75
ptsI	Putative phosphoenol pyruvate:sugar phosphotransferase system enzyme I	836,71	2304,11	0,363	2,75
hisH	Imidazole glycerol phosphate synthase subunit HisH	366,72	1006,24	0,364	2,74
Cp1002_1718	Conserved hypothetical protein	53,57	146,86	0,365	2,74
Cp1002_1194	Hypothetical protein	150,7	401,37	0,375	2,66
rimM	16S rRNA-processing protein	85,69	227,91	0,376	2,66
Cp1002_1976	Conserved hypothetical protein	44,06	115,62	0,381	2,62
Cp1002_0633	Putative inositol monophosphatase family protein	111,67	292,64	0,382	2,62
Cp1002_1495	Conserved hypothetical protein	198,44	519,54	0,382	2,62
Cp1002_0228	DNA polymerase III subunit delta	39,94	103,18	0,387	2,58
Cp1002_1807	Hypothetical protein	867,36	2213,89	0,392	2,55
Cp1002_1553	Hypothetical protein	94,88	237,56	0,399	2,50
entD	4'-phosphopantetheinyl transferase entD	155,05	385,62	0,402	2,49
arsC	Arsenate reductase	61,03	145,17	0,420	2,38
rnpA	Ribonuclease P protein component	1976,43	4675,43	0,423	2,37
groES	Co-chaperonin GroES	4538,76	10603,88	0,428	2,34
plsC	Putative 1-acyl-sn-glycerol-3-phosphate acyltransferase	481,35	1120,16	0,430	2,33
Cp1002_1000	Putative secreted protein	93,94	216,85	0,433	2,31
Cp1002_0266	putative DNA-binding (excisionase) protein	328,97	751,38	0,438	2,28
Cp1002_2011	Putative integral membrane protein	2881,3	6575	0,438	2,28
Cp1002_0397	Conserved hypothetical protein	74,02	166,02	0,446	2,24
carD	Transcriptional regulator	126,15	280,62	0,450	2,22
ppm1	Putative polyprenol phosphate mannosyl transferase1	429,58	945,54	0,454	2,20
glyS	Glycyl-tRNA synthetase	442,66	971,65	0,456	2,20
bglB	Cryptic beta-glucoside bgl operon antiterminator	196,48	430,9	0,456	2,19
gatC	aspartyl/glutamyl-tRNA amidotransferase subunit C	88,14	192,58	0,458	2,18
tyrS	tyrosyl-tRNA synthetase	306,24	662,42	0,462	2,16
dcuA	Anaerobic C4-dicarboxylate transporter	130,42	278,86	0,468	2,14
Cp1002_0643	metalloendopeptidase-like membrane protein	91,49	195,19	0,469	2,13
sepF	Putative cell division protein SepF	668,22	1425,12	0,469	2,13

Cp1002_0548	Conserved hypothetical protein	305,1	650,61	0,469	2,13
Cp1002_1442	Putative membrane protein	742,91	1583,77	0,469	2,13
ftsZ	Cell division protein FtsZ	490,62	1031,88	0,475	2,10
Cp1002_1237	Conserved hypothetical protein	256,06	537,23	0,477	2,10
aroB	3-dehydroquinate synthase	111,97	234,9	0,477	2,10
Cp1002_0942	putative membrane protein	123,18	257,08	0,479	2,09
lpqE	Lipoprotein LpqE	180,89	371,44	0,487	2,05
nrdF	Ribonucleoside-diphosphate reductase beta chain	226,99	463,32	0,490	2,04
ndk1	Nucleoside diphosphate kinase	173,52	349,97	0,496	2,02
yqgF	Putative Holliday junction resolvase	224,48	452,75	0,496	2,02
Cp1002_1165	Conserved hypothetical protein	195,74	394,77	0,496	2,02
ctaD	Cytochrome C oxidase polypeptide I	316,32	633,24	0,500	2,00
Cp1002_0853	Putative thioredoxin-like protein	152,78	305,69	0,500	2,00

RPKM pH- valor de expressão do gene no meio ácido. RPKM controle-valor de expressão do gene no meio controle. *Fold-change* – razão do valor de RPKM pH/RPKM controle. *Fold-change* controle/estresse- razão da expressão entre a condição controle e estresse ácido, somente para facilitar a visualização da repressão do gene nesta condição.

5.9.1. Processo Biológico do estimulon dos genes induzidos na acidez

No estresse ácido observa-se maior número de genes compondo processos (Figura 19) envolvidos no processo da infecção, como no processo de adesão celular o qual 4 genes fazem parte (Cp1002_0023= *secreted protein with 1 motif*; Cp1002_1083 =*secreted protein with 1 motif*; Cp1002_1964 =*secreted protein with 1 motif* e Cp1002_0439= *ABC superfamily ATP binding cassette substrate-binding protein*. Contudo, processo envolvido na regulação da transcrição é destaque entre os 20 primeiros processos. As proteínas regulatórias e sigmas compõem este processo (anexo 11), demonstrando uma regulação imediata de uma resposta específica neste ambiente, presumindo genes relacionados à virulência, necessários à sobrevivência e manutenção da bactéria em um meio tão precário.

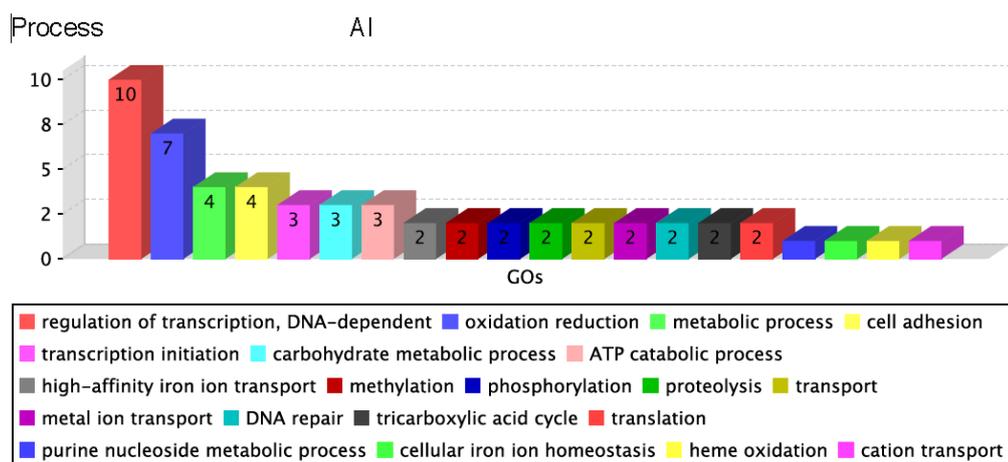


FIGURA 19 - Frequências de genes induzidos pertencentes ao estímulo ácido. Figura obtida pelo programa CSI.

5.9.2. Processo Biológico do estímulo dos genes reprimidos na acidez

Na situação de estresse, como a célula desacelera seu crescimento, é aceitável que a produção de espécie oxigênio reativa seja menor, e assim, o processo de oxido-redução está entre os processos mais reprimidos (figura 20) e a comprovação desta redução pode ser observada no processo transporte onde o gene Cp1002_1670 codificante de *cytochrome c oxidase subunit I* está presente (anexo 12) demonstrando que a regulação da produção aeróbica de energia foi reduzida, confirmando assim, redução no crescimento. Outros processos em destaque foram o de dobramento de proteínas, regulação da transcrição e transporte.

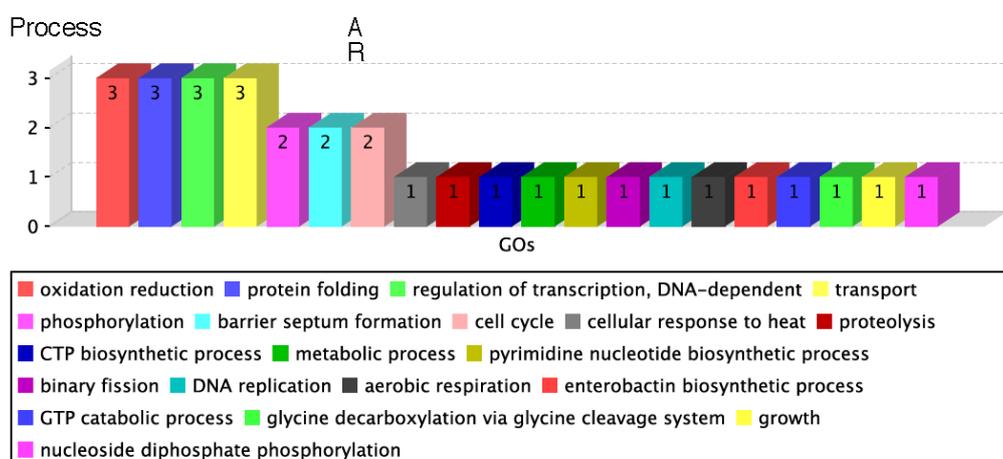


FIGURA 20 - Frequências de genes reprimidos pertencentes ao estímulo ácido. Figura obtida pelo programa CSI.

5.10. Estimulon do estresse térmico, com *fold-change* ≥ 2 ou $\leq 0,5$

O estresse térmico foi a condição que menos representou genes, tanto induzidos quanto reprimidos, quando realizou o corte pelo *fold-change* maior ou igual a dois. Os genes possivelmente apresentam uma expressão maior em fase mais tardia de crescimento, talvez por apresentar maior estabilidade.

Vinte e seis genes foram considerados induzidos neste ambiente com *fold-change* ≥ 2 e seis reprimidos, com *fold-change* $\leq 0,5$. Estes genes estão listados nas tabelas 14 (induzidos) e 15 (reprimidos). O gene *hspR* apresentou um valor de RPKM bem acima do controle, o que nos chamou a atenção e na figura 21 visualiza-se a cobertura dos transcritos no meio térmico e no controle, obtida através do arquivo de saída do Bioscope (.bam).

TABELA 14 - Genes induzidos na condição de estresse térmico, com *fold-change* ≥ 2 .

Genes	Produto	RPKM 50°C	RPKM controle	<i>Fold change</i>
Cp1002_0050	Hypothetical protein	10,97	0	10,97
Cp1002_2046	Conserved hypothetical protein	13,24	1,04	12,73
Cp1002_0622	Putative oxygenase	11,88	1,12	10,61
Cp1002_1060	Putative ABC transporter ATP-binding protein	15,77	2,49	6,33
<i>betA</i>	choline dehydrogenase	209,69	42,02	4,99
<i>yceI</i>	Putative YceI family protein	23,99	5,16	4,65
<i>hspR</i>	Putative chaperone locus transcriptional regulator	1094,19	243,76	4,49
Cp1002_1447	Hypothetical protein	19,19	5,18	3,70
<i>betT</i>	High-affinity choline transport protein	88,45	25,34	3,49
<i>gbsA</i>	Putative betaine aldehyde dehydrogenase	208,35	66,88	3,12
Cp1002_0014	Conserved hypothetical protein	313,12	104,1	3,01
Cp1002_2047	ABC-2 type transporter family protein	28,76	10,02	2,87
<i>gluA</i>	glutamate ABC transport system, ATP-binding protein	34,89	12,91	2,70
Cp1002_0554	Abi_2 superfamily	22,7	8,46	2,68
Cp1002_0061	Conserved hypothetical protein	22,58	8,54	2,64
<i>pepC</i>	Aminopeptidase C	54,83	22,16	2,47
Cp1002_2076	Hypothetical protein	244,69	99,81	2,45
Cp1002_0267	Hypothetical protein	2457,43	1031,26	2,38
Cp1002_0007	Conserved hypothetical protein	295,15	123,9	2,38
Cp1002_0568	L-asparagine permease	34,78	15,12	2,30
<i>ahpD</i>	Alkyl hydroperoxide reductase AhpD	1115,29	523,66	2,13
<i>pgm</i>	Phosphoglucomutase	71,35	33,57	2,13
Cp1002_1094	Putative membrane protein	87,69	42,14	2,08
Cp1002_1013	Putative secreted protein - YceI	244,25	118,62	2,06
Cp1002_1221	Putative membrane protein	335,15	165,07	2,03
<i>frr</i>	ribosome recycling factor	1732,38	858,55	2,02

RPKM 50°C- valor de expressão do gene no meio térmico. RPKM controle-valor de expressão do gene no meio controle. *Fold-change* – razão do valor de RPKM 50°C/RPKM controle.

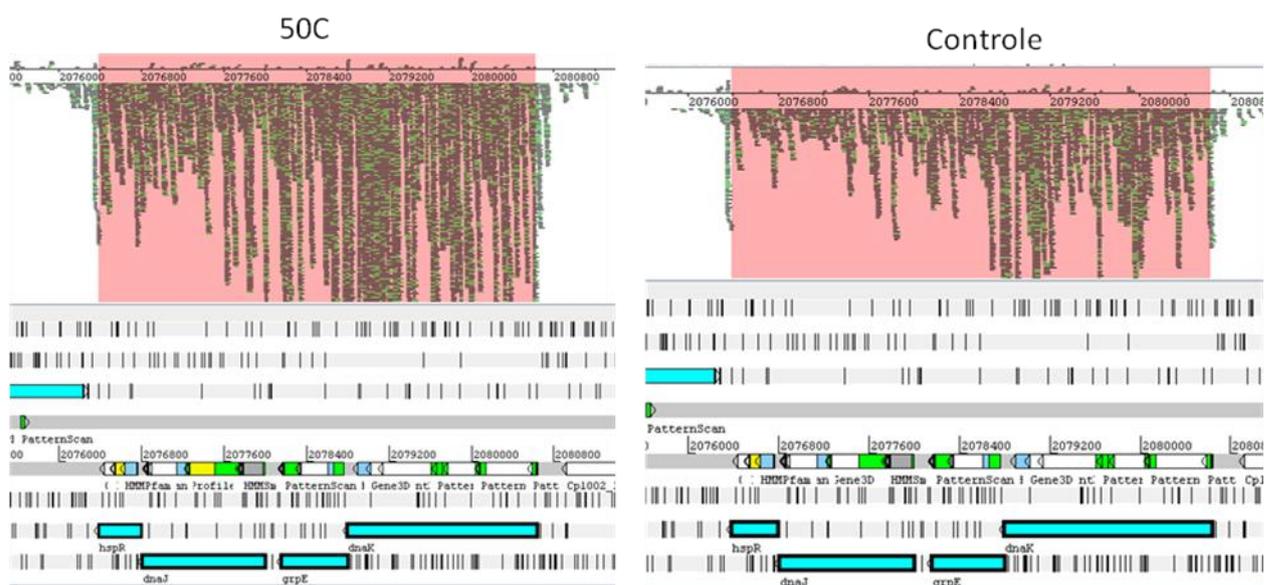


FIGURA 21 - Cobertura dos genes (operon) *dnaK-grpE-dnaJ-hspR*, envolvidos na resposta ao choque térmico. Genes expressos no meio sob estresse térmico à esquerda e no controle, à direita. Observa-se a diferença na expressão entre as duas condições. Obtida pelo Artemis através do arquivo de saída .bam do Bioscope.

Entre os reprimidos, 50% dos genes expressos diferencialmente com *fold-change* $\leq 0,5$ codificam proteínas hipotéticas, como se observa na tabela 15.

TABELA 15 - Genes reprimidos no estresse térmico, com *fold-change* $\leq 0,5$.

Gene	Produto	RPKM 50°C	RPKM controle	<i>fold-change</i>	<i>fold-change</i> controle/estresse
Cp1002_1507	Hypothetical protein	18,37	57,91	0,317	3,152
Cp1002_1539	Hypothetical protein	50,78	145,66	0,348	2,868
<i>argB</i>	acetylglutamate kinase	69,3	183,99	0,376	2,655
Cp1002_1356	Hypothetical protein	47,98	117,42	0,408	2,447
<i>atpC</i>	FOF1 ATP synthase subunit epsilon	5237,38	11411,9	0,458	2,179
<i>pflAC</i>	Formate acetyltransferase	1144,63	2363,39	0,484	2,064

RPKM 50°C- valor de expressão do gene no meio térmico. RPKM controle-valor de expressão do gene no meio controle. *Fold-change* – razão do valor de RPKM 50°C/RPKM controle. *Fold-change* controle/estresse- razão da expressão entre a condição controle e estresse ácido, somente para facilitar a visualização da repressão do gene nesta condição.

5.10.1. Processo Biológico do estímulo dos genes induzidos no choque térmico

No estresse térmico, observamos que o processo de oxido-redução está entre os processos mais representados, seguido do processo catabólico (Figura 22). Através do processo

catabólico, é possível produzir mais energia para direcionar à síntese de proteínas específicas à sobrevivência neste ambiente. Os genes que participam de cada processo podem ser visualizados no anexo 13.

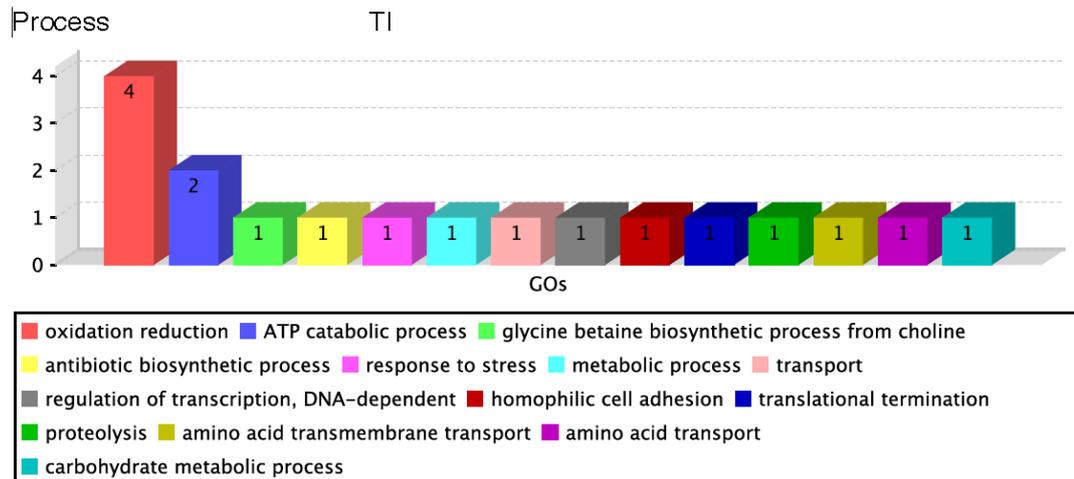


FIGURA 22 - Frequências de genes induzidos pertencentes ao estímulo térmico. Figura obtida pelo programa CSI.

5.10.2. Processo Biológico do estímulo dos genes reprimidos no choque térmico

Poucos genes foram selecionados com o corte definido nesta condição, mas se observa na figura 23 que são genes relacionados a biossíntese e metabolismo, justificando a redução no crescimento da bactéria no ambiente. No anexo 14 visualiza-se o gene de cada processo.

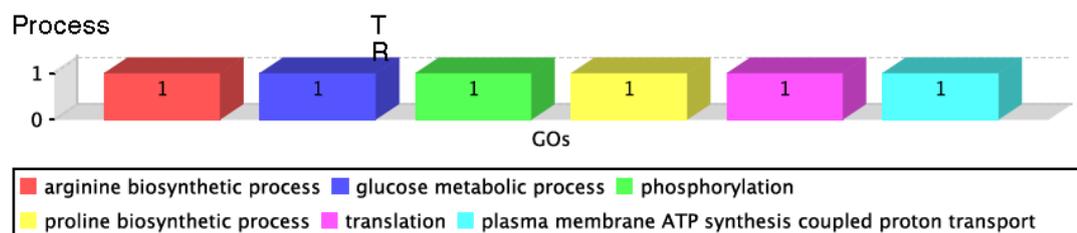


FIGURA 23 - Frequências de genes reprimidos pertencentes ao estímulo térmico. Figura obtida pelo programa CSI.

5.11. Expressão dos Genes codificantes de fatores sigmas

Os sigmas são importantes reguladores da expressão dos genes, por isso, é de extrema importância observar quais sigmas estão sendo induzidos ou reprimidos em cada perturbação ambiental. O genoma de *C. pseudotuberculosis* linhagem 1002 contém sete sigmas anotados e todos estão listados na tabela 16, com os respectivos valores de expressão em cada condição. Na figura 24, observamos o *heat map*, que é composto por um código de cores a fim de facilitar a visualização dos níveis de expressão gênica, entre estresse e controle.

TABELA 16 - Expressão dos genes codificantes dos fatores sigmas.

Gene	Produto	Controle	2M	50°C	pH	fold-change/2M	fold-change/50°C	fold-change/pH
<i>sigA</i>	RNA polymerase sigma factor rpoD (Sigma-A)	259,94	582,09	410,5	252,56	2,2393	1,5792	0,9716
<i>sigB</i>	RNA polymerase sigma factor sigB	384,38	252,65*	359,98*	1187,13	0,6573	0,9365	3,0884
<i>sigC</i>	RNA polymerase sigma factor	153,47	80,03*	135,77*	272,8	0,5215	0,8847	1,7775
<i>sigD</i>	RNA polymerase sigma factor SigD	340,83	534,59	461,76	299,61*	1,5685	1,3548	0,8791
<i>sigE</i>	RNA polymerase sigma factor SigE	283,22	159,65*	277,11*	729,05	0,5637	0,9784	2,5741
<i>sigH</i>	RNA polymerase sigma factor RpoE	320	598,88	595,45	1175,3	1,8715	1,8608	3,6728
<i>sigK</i>	ECF family sigma factor K	46,03	56,19*	72,55	47,84*	1,2207	1,5761	1,0393
<i>sigM</i>	ECF family sigma factor M	3,47	17,25	12,86*	18,08	4,9712	3,7061	5,2104

2M-estresse osmótico; 50°C-estresse térmico; pH estresse ácido. * - genes que apresentaram valor de *p-value* > 0,001; não foram significativo.

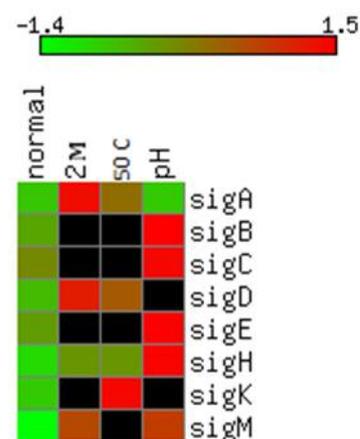


FIGURA 24 - Mapa de cores relacionado ao nível de expressão dos genes codificantes dos fatores sigmas. Célula de cor preta indica o gene que não foi considerado diferencialmente expresso na condição, apresentando p -value > 0,001. Cada célula colorida representa o valor de expressão do gene em cada condição. A maior expressão está em vermelho e a menor expressão está em verde. O valor intermediário está como marrom. A escala mostra o valor em \log_2 fold-change. Imagem criada no matrix2png (<http://www.bioinformatics.ubc.ca/matrix2png/>)

5.12. Ilhas de Patogenicidade

Para selecionar os genes presentes nas ilhas de patogenicidade (preditos in silico-Soares Dissertação-2009) fez-se o corte baseando nos resultados do DEGseq. Considerou-se os genes que foram significativos (p -value < 0,001) independentes do valor de $fold$ -change. Na tabela 17 observam-se estes resultados. Na figura 25, foi construído o mapa de cores representando o nível de expressão destes genes. Pela figura podem-se perceber genes que não foram considerados diferencialmente expressos (cor preta) em uma situação, mas sim em outra (cores diferentes).

TABELA 17 - Genes presentes nas ilhas de patogenicidade que foram considerados diferencialmente expressos.

Gene	RPKM Controle	RPKM 2M	RPKM 50°C	RPKM pH	Produto	fold-change 2M	fold-change 50°C	fold-change pH
<i>cat</i>	305,98	203,76*	309,75	1332,97	Catalase	0,67	1,01	4,36
Cp1002_0023	3,27	9,88*	8,64*	24,32	Conserved hypothetical protein	3,02	2,64	7,44
Cp1002_0025	26,14	155,39	56,92	131,15	CRISPR-associated protein	5,94	2,18	5,02
Cp1002_0048	1098,15	512,16	799,31*	2133,44	Conserved hypothetical protein	0,47	0,73	1,94
Cp1002_0050	0	2,24	10,97	0	Hypothetical protein	2,24	10,97	0
Cp1002_0054	122,88	254,2	142,05	145,94	Hypothetical protein	2,07	1,16	1,19

Cp1002_0174	190,54	43,1	76,77	44,99	Hypothetical protein	0,23	0,40	0,24
Cp1002_0175	112,88	347,98	145,07	78,64*	Hypothetical protein	3,08	1,29	0,70
Cp1002_0176	39,45	247,57	69,53	106,93	Hypothetical protein	6,28	1,76	2,71
Cp1002_0177	4,03	34,83	8,53*	0*	Hypothetical protein	8,64	2,12	0,00
Cp1002_0178	50,68	81,18	32,85*	55,93*	Hypothetical protein	1,60	0,65	1,10
Cp1002_0179	56,89	105,62	66,78*	111,13	Hypothetical protein	1,86	1,17	1,95
Cp1002_0980	87,2	131,44	122,75	68,86*	Putative membrane protein	1,51	1,41	0,79
Cp1002_0988	99,82	711,84	189,65	170,41	Putative secreted protein with LPxTG motif	7,13	1,90	1,71
Cp1002_0989	61,25	177,26	99,3	27,61*	Putative surface-anchored membrane protein	2,89	1,62	0,45
Cp1002_0990	6,12	19,37	6,47*	18,2	Putative integrase	3,17	1,06	2,97
Cp1002_0991	79,74	168,59	93,05*	346,27	Hypothetical protein	2,11	1,17	4,34
Cp1002_1447	5,18	4,48*	19,19	7,71*	Hypothetical protein	0,86	3,70	1,49
Cp1002_1454	2,39	30,93	10,1*	85,24	Hypothetical protein	12,94	4,23	35,67
Cp1002_1457	20,72	49,38	27,46*	39,17	Putative methyltransferase	2,38	1,33	1,89
Cp1002_1462	140,79	191,84	135,66*	74,46*	Hypothetical protein	1,36	0,96	0,53
Cp1002_1463	14,63	25,28*	9,29*	55,15	Hypothetical protein	1,73	0,63	3,77
Cp1002_1467	82,59	30,37	73,87*	71,61*	Putative cell-surface hemin receptor	0,37	0,89	0,87
Cp1002_1469	138,19	275,07	187,9	358,77	Inosine 5-monophosphate dehydrogenase	1,99	1,36	2,60
Cp1002_1471	510,7	803,73	986,98	1529,58	Conserved hypothetical protein	1,57	1,93	3,00
Cp1002_1552	127,56	139,01*	160,2	122,35*	ABC transport system ATP-binding protein	1,09	1,26	0,96
Cp1002_1553	237,56	131,07*	125,62	94,88	Hypothetical protein	0,55	0,53	0,40
Cp1002_1556	62,13	10,66	22,19*	16,73	Membrane protein, YeeE/YedE family	0,17	0,36	0,27
Cp1002_1558	4,12	14,25*	13,09*	32,72	Hypothetical protein	3,46	3,18	7,94
Cp1002_1560	8,11	17,5*	15,17*	21,33	Putative transport system membrane protein	2,16	1,87	2,63
Cp1002_1565	30,24	56,99	46,52*	60,89	Hypothetical protein	1,88	1,54	2,01
Cp1002_1910	100,26	81,58*	148,97	264,59	Conserved hypothetical protein	0,81	1,49	2,64
Cp1002_1914	14,54	41,85	29,61*	41,12	ABC superfamily ATP binding cassette transporter	2,88	2,04	2,83
Cp1002_1928	78,47	122,83	94,23*	54,31*	Putative integral membrane protein	1,57	1,20	0,69
Cp1002_1929	38,05	96,05	34,05*	203,17	Conserved hypothetical protein	2,52	0,89	5,34
dcd	736,43	582,09*	652,91*	391,32	Deoxycytidine triphosphate deaminase	0,79	0,89	0,53
fagA	44,86	26,41*	71,16	31,85*	integral membrane protein -FagA	0,59	1,59	0,71
fagD	1380,39	1001,02*	1853,52	3707,7	iron siderophore binding protein-FagD	0,73	1,34	2,69
fetA	37,75	41,75*	60,24	46,63*	Putative iron ABC transport system	1,11	1,60	1,24
fetB	77,53	55,8*	102,79	40,67	Putative iron transport system membrane protein	0,72	1,33	0,52
fetS	230,72	134,89*	267,31	169,03*	Putative iron transport system binding	0,58	1,16	0,73

					(secreted) protein			
frgA	16,7	49,79	31,59*	32,65	Putative siderophore biosynthesis related protein	2,98	1,89	1,96
glpT	31,15	28,25*	23,47*	52,13	Glycerol-3-phosphate transporter	0,91	0,75	1,67
lcoS	83,11	113,65	77,64*	52,2*	Two-component sensor protein, sensor histidine kinase	1,37	0,93	0,63
mfsD1	25,61	57,2	49,5	36,34*	Major facilitator superfamily domain-containing protein 1	2,23	1,93	1,42
mgtE	6,69	8,5*	11,66*	29,28	Mg2 transporter mgtE	1,27	1,74	4,38
mnhG	54,62	56,3*	72,67*	71,63	Putative H anti-porter subunit G	1,03	1,33	1,31
phoB	47,31	105,15	52,49*	54,6*	Two-component regulatory protein	2,22	1,11	1,15
phuC	4,94	20,12	20,91	23,11	Iron(III) dicitrate transport permease-like protein yusV	4,07	4,23	4,68
pld	2,36	1,53*	0*	13,44	Phospholipase D precursor	0,65	0	5,69
sigK	46,03	56,19*	72,55	47,84*	ECF family sigma factor K	1,22	1,58	1,04
udg	39,12	113,22	75,71	73,47	UDP-glucose 6-dehydrogenase	2,89	1,94	1,88
ureA	16,17	15,52*	13,3*	39,2	Urease gamma subunit	0,96	0,82	2,42
ureD	141,36	169,21	172,44	122,58*	Urease accessory protein UreD	1,20	1,22	0,87
ureE	198,59	187,27*	243,81	149,39*	Urease accessory protein UreE	0,94	1,23	0,75
ureF	98,29	115,58	107,95*	77,23*	Urease accessory protein UreF	1,18	1,10	0,79
ureG	147,83	220,17	205,02	261,59	Urease accessory protein UreG	1,49	1,39	1,77

2M-estresse osmótico; 50°C- estresse térmico; pH-estresse ácido. *Fold-change* maior que 1= induzido; *fold-change* menor que 1= reprimido.

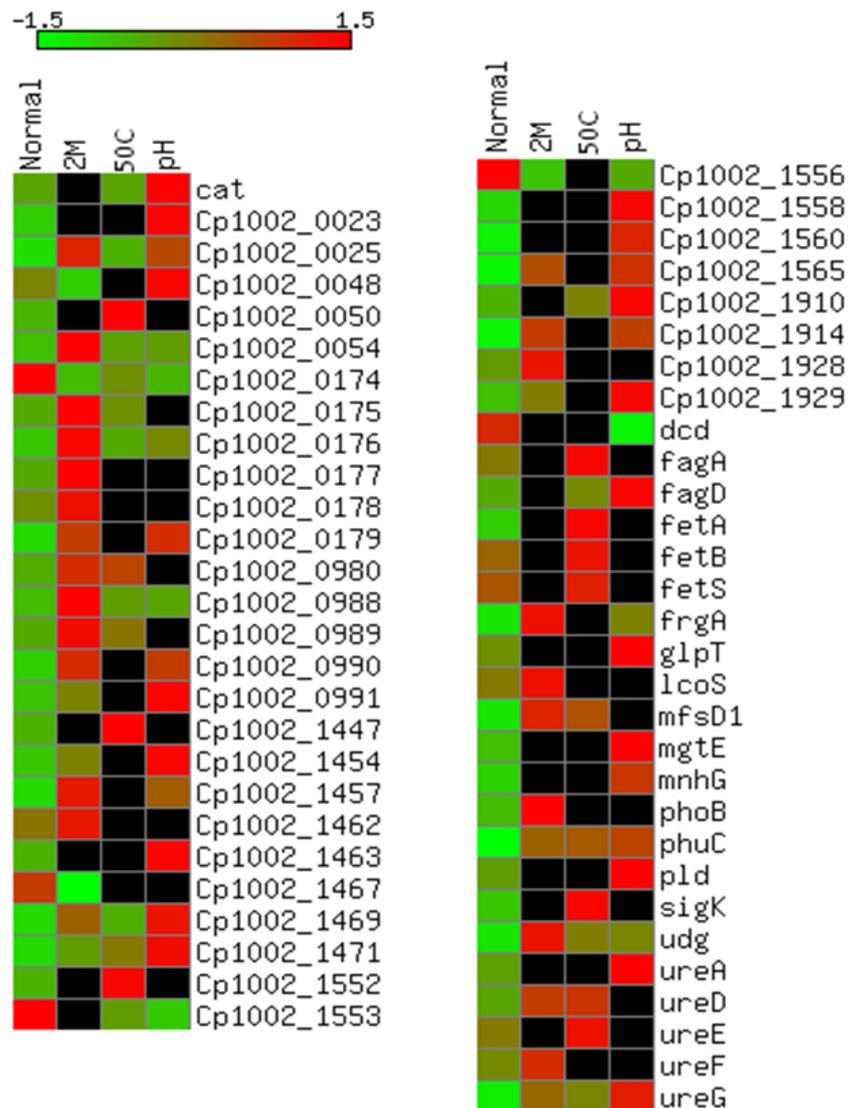


FIGURA 25 - Genes diferencialmente expressos presentes nas ilhas de patogenicidade de *C.pseudotuberculosis* 1002.

Nas colunas estão as condições e nas linhas estão os genes. Células de cor preta representam genes que não entraram na análise na condição em questão, pois não foram considerados diferencialmente expressos. Não foram selecionados naquela condição. Vermelhos são genes induzidos e verde genes reprimidos. Marron, ou verde escuro são genes com expressão intermediária. Escala representada pelo valor de log₂ fold-change. Imagem criada no matrix2png (<http://www.bioinformatics.ubc.ca/matrix2png>).

5.13. Produtos gênicos confirmados, *in silico* ou experimental, por outros trabalhos do grupo LGCM, envolvendo *C.pseudotuberculosis* 1002

Muitos trabalhos envolvendo a *C.pseudotuberculosis* 1002 estão sendo realizados por nosso grupo no intuito de desvendar moléculas que estejam relacionadas com a virulência, patogenicidade, defesa, para buscar soluções eficazes contra este microrganismo patogênico. Assim, os resultados de diferentes trabalhos (proteômica, análise *in silico*-SurfG⁺, imunoproteômica, *phage display*) foram cruzados e foi observado, dentre estes, quais apresentaram um valor de expressão significativo em cada condição. Somente os genes que apresentaram *fold-change* ≥ 2 foram analisados. A lista com todos os genes averiguados estão no anexo 15. O objetivo foi conhecer o nível de expressão destes genes sob alteração ambiental e no início da fase exponencial.

6. DISCUSSÃO

Este trabalho pioneiro no Brasil teve como objetivo detectar genes diferencialmente expressos em diferentes condições ambientais de *C.pseudotuberculosis* 1002. Porém, além do próprio ambiente influenciando/direcionando a expressão dos genes, em procariotos diversos fatores regulam a expressão gênica e a abundância de proteínas em vários níveis: iniciação da transcrição; força do promotor; configuração do promotor; fatores de transcrição; terminação da transcrição; estabilidade e taxa de rotatividade do mRNA; uso do códon; iniciação da tradução e alongação; e dobramento das proteínas, degradação e localização celular (Karlin e Mrázek, 2000).

O sequenciamento do RNA é uma tecnologia atrativa para estudos da expressão gênica nos organismos procariotos possibilitando uma aproximação da realidade do ambiente celular em diversas condições investigadas. Assim, *C. pseudotuberculosis* linhagem 1002 foi alvo de nossos estudos em situações que mimetizam alguns ambientes encontrados por esta bactéria patogênica no processo da infecção. Estresse osmótico, térmico e acidez foram simulados para desvendarmos os genes envolvidos na resistência desta bactéria nestas perturbações ambientais. Para selecionarmos genes importantes para a virulência, ou envolvidos com a resistência da bactéria em cada situação, foi necessário identificar os genes diferencialmente expressos.

A quantidade de leituras e cobertura obtida (Tabela 2) em cada situação nos permitiu identificar estes genes que podem se tornar principais candidatos ao desenvolvimento de uma vacina eficaz contra LC. No trabalho com RNAseq de Camarena et al (2010) utilizando Illumina, conseguiram 25,3 vezes de cobertura no genoma de *Acinetobacter baumannii*, combinando dois sequenciamentos que totalizaram 3.596.474 leituras únicas e tentaram melhorar com três ciclos de depleção de rRNA, mas não obtiveram sucesso. Isabella e Clark (2011) utilizando a plataforma SOLiD™ para obtenção dos transcritos de *Neisseria gonorrhoeae*, analisando 4 condições diferentes, obtiveram uma média de 1.500.000 leituras mapeadas. Neste trabalho, em apenas um sequenciamento, foi possível obter uma média de 1.860.000 leituras mapeadas, o que resultou em uma média de 46,5 vezes de cobertura.

A confiabilidade nos resultados é de extrema importância para garantir a escolha certa de um alvo para estudos futuros envolvendo diversas pesquisas sem perder tempo ou investimento. O programa BioScope utilizado para a obtenção dos dados garantiu a confiabilidade no momento em que tratou os dados, através da normalização.

A experiência com os dados de microarranjos tem mostrado que a normalização é uma etapa crítica na execução do *pipeline*, permitindo estimativa precisa e a detecção da expressão

diferencial. O objetivo da normalização é remover os efeitos técnicos que ocorrem nos dados, para assegurar que estes, tenham o mínimo de impacto nos resultados. No entanto, o procedimento para gerar dados do RNAseq é diferente dos dados de microarranjos, e assim, a normalização não é aplicada exatamente da mesma maneira. No RNAseq padroniza-se pelo número de leituras no experimento. A pesquisa pode diferir nas hipóteses de distribuição para inferir a diferença, mas o consenso é utilizar o número total de leituras no experimento (Robinson e Oshlack, 2010).

No programa Bioscope a normalização é ajustada pela medida RPKM, segundo (Mortazavi et al., 2008), em que a contagem das leituras ocorre por Reads per Kilobase per Million Mapped (RPKM). Neste caso, a medida de densidade reflete a concentração molar de um transcrito, pela normalização do tamanho do gene e o número total de leitura em cada gene. Para as análises dos genes diferencialmente expressos utilizou-se o programa DEGseq e para selecionarmos os genes após o corte definido no DEGseq, foi desenvolvido o programa *DBtransProt* para obtenção dos dados de maneira mais rápida e confiável.

6.1. Análise da expressão dos genes mais caracterizados em *C. pseudotuberculosis*

Os genes mais caracterizados de *C. pseudotuberculosis* linhagem 1002 foram analisados mesmo quando não apresentaram uma diferença significativa na expressão em relação ao controle. Além do *pld*, e *genes fag*, outros genes já mostraram essenciais para o estabelecimento de LC causada por *C.pseudotuberculosis*, como gene codificante da serina protease e componentes da parede celular.

6.1.1. Fosfolipase D

O gene *pld*, que codifica a proteína fosfolipase D (Pld), uma esfingomielinase implicada na virulência de *C. pseudotuberculosis* (Mcnamara et al., 1995), apresentou pequeno número de leituras cobrindo a região transcricionalmente ativa e não foi classificado como gene diferencialmente expresso em duas condições (osmótico e temperatura) (Tabela 3). Somente na condição de acidez, apresentou um *p-value* significativo (*p-value* < 0,001) e foi considerado induzido. Segundo (Mckean et al., 2007a), o gene *pld* é regulado por dois fatores: térmico e densidade celular. Eles conseguiram demonstrar, trabalhando com *C. pseudotuberculosis* linhagem C231, que a expressão de *pld* em todas as densidades investigadas, foi reprimida pelo choque térmico (43⁰C) (Mckean et al., 2007b). Este estresse foi capaz de reduzir a expressão de *pld* em nível similar ao basal. O que indica que o

mecanismo de termoregulação predomina em relação à densidade. O experimento mostrou que mesmo em uma cultura de densidade baixa (DO=0,1), onde a expressão de *pld* não é alta, o nível pode ser ainda mais reduzido, quando o estresse térmico está presente. Neste trabalho com a linhagem 1002, o gene *pld* no estresse térmico apresentou uma expressão similar ao controle e foi considerado reprimido (Tabela 3).

No caso do estresse ácido, cujo objetivo foi simular o ambiente do macrófago, o resultado obtido nos aproxima daqueles alcançados por estes mesmos autores (McKean et al., 2007a). O experimento demonstrou que o gene *pld* foi altamente expresso dentro de macrófagos, porém, *in vitro*, existe dependência da densidade celular. Devido a propriedade desta enzima, é provável que a Pld intracelular afete a função e viabilidade dos macrófagos, seja pela destruição de membranas nestas células (McKean, Davies e Moore, 2007a) ou por interromper as rotas sinalizadoras do hospedeiro (Schwarzer et al., 1998). Na figura 13, podemos visualizar que não ocorreu uma alta expressão do gene no estresse por acidez. A cobertura está baixa, mas se observa uma expressão maior que nos outros estresses, com *fold-change* de quase 6 vezes a mais que o controle (Tabela 3), possivelmente devido à regulação dependente da densidade celular. Em experimento realizado em macrófagos, esta dependência pela densidade celular na expressão do gene não foi observada (McKean et al., 2007a).

Pacheco et al (2011) desenvolveram um trabalho visando o exoproteoma comparativo entre duas linhagens de *C. pseudotuberculosis* e não encontraram a proteína codificada por este gene, na linhagem 1002, mas vale ressaltar que o trabalho foi realizado em outra densidade (DO=0,4) e em outro meio de cultivo (meio mínimo). Este resultado leva a associar o gene *pld* ao grau de virulência do organismo, uma vez que, *C.pseudotuberculosis* 1002 vem sendo tratada como cepa naturalmente de baixa virulência, como já relatado (Ribeiro et al., 1991) (Meyer et al., 2002) e, no trabalho de Pacheco et al (2011) na cepa C231, que apresenta maior grau de virulência, foi localizado a proteína Pld.

Estudo com *C.pseudotuberculosis* em que a Pld foi inativada, demonstrou convincentemente a necessidade da Pld para estabelecimento de LC, e persistência da bactéria no hospedeiro (Mcnamara et al., 1994). A linhagem 1002 foi testada como vacina viva atenuada contra LC em cabras e a administração desta bactéria não causou lesão no animal (Meyer et al., 2002), porém, em camundongos causou danos (Pacheco, comunicação pessoal). No trabalho realizado por Pacheco et al (2011), foi realizada uma busca, *in silico*, pelo promotor do gene *pld* e foi observada uma mutação por transversão, de timina para guanina, na suposta região -35 do promotor do gene *pld*. Um reconhecimento pela sequência específica do

promotor é de suma importância e crítica para uma regulação da transcrição bem sucedida (Song et al., 2008). Desta forma, pode-se sugerir que a baixa expressão deste gene tenha sido influenciada por esta alteração no promotor. Mas ainda assim, a expressão foi maior na acidez e neste caso, pode-se sugerir que um regulador ativo apenas na condição da acidez, tenha contribuído para o aumento na expressão do gene.

6.1.2. Genes *fag*

A bactéria e outros patógenos produzem uma ampla faixa de determinantes de virulência requeridos para a patogenicidade, muitos dos quais são regulados por fatores ambientais. O baixo nível do ferro tem mostrado induzir a expressão de diferentes toxinas bacterianas e fatores de virulência (Sritharan, 2006).

Alguns genes podem ser regulados por mais de um fator ambiental, como o gene invasiva de *Yersinia pseudotuberculosis*, que é regulado em resposta a uma variedade de sinais ambientais, como temperatura, fase de crescimento, nutriente, osmolaridade e pH (Nagel et al., 2001). Os genes *fagB* e *fagC* em *C.pseudotuberculosis* 1002 não foram considerados diferencialmente expresso (p-value > 0,001) em nenhuma condição testada (Tabela 3).

Segundo Billington et al (2002), a atenuação do gene *fagB* não mostrou nenhum defeito na utilização do ferro *in vitro*, mas o mutante mostrou redução na virulência de LC em um modelo de bode quando comparado a linhagem selvagem (controle). A regulação do operon *fagABC* e do gene *fagD* supostamente ocorre pelo ferro, mas é possível que necessitem de outro, ou de outros fatores, *in vivo*, para alcançar indução completa. Pouco se sabe sobre o uso das fontes de ferro por *C. pseudotuberculosis*, mas tudo indica que possa adquirir ferro a partir de uma variedade de fonte.

O operon *fagABC*, que codifica proteínas envolvidas no sistema de absorção do ferro, mostrou uma alternância na expressão nas diferentes condições apresentadas. O operon está relacionado ao sistema de transporte de ferro, mas associado a membrana citoplasmática, apresentando proteínas hidrofóbicas (Billington et al., 2002). Na tabela 3 pode-se observar os valores da expressão do operon em cada condição testada, juntamente com o valor *fold_change*.

Na condição de estresse por acidez e osmótico o operon *fagABC* não foi considerado diferencialmente expresso. No entanto, no estresse onde foi aplicado a temperatura de 50 graus, o gene *fagA* foi induzido e considerado diferencialmente expresso e os dois outros genes componentes deste operon, não foram considerados diferencialmente expressos. Na tabela 03

pode-se observar os valores de RPKM para cada gene. Esta alternância de transcrição nos aproxima dos resultados obtidos por (Güell et al., 2009) no trabalho com *Mycobacterium pneumoniae* cuja expressão surpreendente foi observada dentro de operons, onde transcrição alternativa foi apresentada em resposta a diferentes condições de crescimento. Segundo este grupo, provavelmente devido a uma regulação por RNA antisense o organismo consiga alterar a estrutura do operon levando a uma transcrição diferente dentro dos operons, visando à sobrevivência em um ambiente hostil. Este resultado acrescentou um nível de complexidade na regulação gênica em bactérias nunca antes imaginadas.

O gene *fagD*, que também codifica proteínas envolvidas no sistema de absorção de ferro (Billington et al., 2002) foi classificado como diferencialmente expresso (*pvalue* < 0,001) nas condições de estresse térmico (valor de RPKM = 1853,52) e estresse ácido (valor de RPKM = 3707,7) e em relação ao controle (RPKM = 1380,39) foi considerado como gene induzido (tabela 03). Acredita-se que o gene *fagD* esteja envolvido também com a virulência em *C.pseudotuberculosis* por fazer parte do sistema de absorção de ferro associado com a parede celular. No estresse osmótico não foi considerado como gene diferencialmente expresso em relação ao controle. Pacheco e colaboradores (2011) também conseguiram identificar a proteína FagD na linhagem 1002, mas não o operon *fagABC*. O gene *fagD* está localizado *downstream* ao operon *fagABC* que por sua vez está localizado na fita reversa *downstream* ao gene *pld*.

O gene *fagD* é regulado pelo *dtxR* (figura 15) um gene cujo produto é responsável pela repressão da toxina dependente de ferro e sideróforo (Schmitt e Holmes, 1991). O gene *dtxR* é regulado pelo ferro, assim, na presença de ferro intracelular suficiente, o complexo DtxR-Fe²⁺ se liga imediatamente *upstream* ao gene *fagD* e inibe sua transcrição. Quando o nível de ferro está baixo, existe uma indução do gene *dtxR*, mas ele não pode se ligar no sitio ligante na ausência do ferro (Sritharan, 2006). Possivelmente, em *C.pseudotuberculosis* 1002, o nível de ferro estava baixo, não houve a formação do complexo DtxR-Fe²⁺, e o gene *fagD* foi altamente induzido. No meio BHI, pela sua composição, pode ser que uma pequena parcela de ferro esteja presente, mas acredita-se que esteja em complexo com grupo heme, pela possível presença de sangue na composição do meio, mas não existe confirmação disso.

Os valores de expressão do gene *dtxR* podem ser visualizados na tabela 4. O gene *dtxR* aumentou sua expressão somente na condição de acidez, mas como não houve a formação do complexo por não haver ferro suficiente, não reprimiu o gene *fagD* e esta talvez seja a causa da maior expressão do gene. Curioso que o gene *dtxR* foi induzido somente na condição de acidez

e foi a condição que mais houve expressão do gene *fagD*. Nas situações onde *dtxR* foi reprimido, a expressão de *fagD* aumentou, mas não tanto. E ainda mais curioso é o fato de no estresse osmótico ter sido a única situação em que *fagD* foi reprimido em relação ao controle e o *dtxR* nesta situação foi reprimido, mas não foi considerado diferencialmente expresso. É bem provável que existam outros fatores envolvidos nesta regulação, por isso novos estudos devem ser realizados envolvendo estes genes para chegar a uma conclusão precisa.

Na condição de estresse térmico, os valores de RPKM do gene *dtxR* entre estresse e controle foi praticamente o mesmo e no entanto foi considerado diferencialmente expresso e os valores de RPKM no estresse osmótico estão bem diferentes, mas não foi considerado diferencialmente expresso. Não sabemos se é um viés do programa e por isso outras análises devem ser realizadas para esclarecer este resultado.

6.1.3. Serina protease

O novo antígeno de 40kDa apresentado por (Walker et al., 1994) e caracterizado por Wilson, Brandon e Walker em 1995, foi apresentado como um tipo de proteína com atividade de serina protease, porém, na pesquisa pelo banco de dados não identificaram qualquer homologia com sitio ativo de serina protease. Então, classificaram a proteína de 40kDa, como uma subclasse de serina protease não caracterizada. Esta proteína foi identificada como uma nova esperança para a caprinovincultura, uma vez que se revelou como um antígeno protetor contra LC (Wilson et al., 1995).

Em *C.pseudotuberculosis* 1002, os dados do transcriptoma confirmaram um gene precursor de serina protease- endoS (CP40), mas com uma expressão praticamente nula, muito similar ao controle, em todas as situações testadas. (Pacheco et al., 2011), também não encontraram esta proteína nos dados de exoproteoma de *C.pseudotuberculosis* 1002.

Analisando todo o transcriptoma de *C.pseudotuberculosis* 1002, encontramos quatro supostas proteínas apresentando a característica de uma serina protease, e entre elas, três foram consideradas diferencialmente expressas nas condições testadas (Tabela 5). São elas *Trypsin family serine protease*, *Putative serine protease*, *Putative trypsin-like serine protease*.

Em procarioto, as proteases exercem um papel significativo nas respostas adaptativas às mudanças no ambiente extracelular, facilitando a reestruturação do seu proteoma. Serina proteases são enzimas com resíduo nucleofílico **Ser** no sítio ativo e acredita-se constituir quase 1/3 de todas as enzimas proteolíticas conhecidas. Incluem exopeptidase e endopeptidase. Em

procariotos atuam em processos associados à sinalização celular, defesa, e desenvolvimento (Tripathi e Sowdhamini, 2008).

A protease extracelular bacteriana dentro do hospedeiro infectado pode exibir uma magnitude de potenciais patogênicos. Sua insensibilidade aos inibidores de proteases presentes no hospedeiro e a habilidade em inativar algumas destas, fazem as proteases acentuadamente nocivas. Estudos com proteases estafilocóccicas mostraram que elas são capazes de interagir com mecanismos de defesa do hospedeiro e componentes do tecido, bem como modificar outros fatores de virulência. O controle é dependente da densidade celular (Dubin, 2002), o que leva a supor que a maioria das proteases de *C.pseudotuberculosis* 1002 não apresentou uma alta expressão, por desenvolver o trabalho em uma densidade baixa.

O gene (Cp1002_0210) codificante da suposta Trypsin family serine protease, apresentou uma expressão maior (*fold-change* = 1,7823) na condição de estresse osmótico. Esta suposta proteína apresenta o domínio Colicin_V, que acredita-se atuar na patogênese e defesa, observada somente em bactéria Gram-positiva (Actinobacteria). O domínio Colicin_V permite eliminar alvos celulares pelo rompimento do seu potencial de membrana e pode contribuir na patogênese e/ou defesa. Talvez o domínio combinatório de tripsina com colicin_V seja requerido à bactéria que vive em condições desafiadoras (Tripathi e Sowdhamini, 2008), como, por exemplo, no estresse osmótico.

A suposta serina protease, cujo gene (Cp1002_0662) apresentou uma expressão duas vezes maior no estresse ácido em relação ao controle (*fold-change* = 2,61), possui o domínio PDZ de uma *trypsin-like serine proteases*, tais como DegP/HtrA, que são proteínas envolvidas na resposta ao choque térmico, função chaperona e apoptose. Considerando ainda que a *trypsin-like serine protease* apresenta diversas outras funções como lise da parede celular, regulação da transcrição, toxinas, enzimas fibrinolíticas, entre outros (Tripathi e Sowdhamini, 2008), é provável que no estresse ácido que simula a condição do macrófago, esta enzima atue para que a bactéria consiga permanecer no macrófago mesmo com tantas condições desfavoráveis. Seja como chaperona ou regulando a transcrição de outros genes, ou como lise da membrana do macrófago.

No estresse térmico, o gene *sprX*, codificando uma *Putative trypsin-like serine protease* foi o mais induzido entre as condições e apresenta domínio Tryp (Tripathi e Sowdhamini, 2008), que está associado a proteólise, além das funções citadas acima. Neste ambiente de alta temperatura, é bem provável que esteja envolvida na resposta ao choque térmico, atuando sua função chaperona.

6.1.4. Componentes da parede celular e modificação lipídica

O principal constituinte da parede celular no grupo micobactéria e gêneros relacionados é o ácido micólico formado por 2-álquil e 3-hidroxi ácido graxo, de alto peso molecular, o qual fornece a primeira linha de defesa contra condições letais do ambiente (Yuan et al., 1995). Cada espécie de micobactéria contém um ácido micólico característico com diferentes funções químicas, incluindo dupla ligação, anéis de ciclopropano, ramificação do grupo metil, que são introduzidos dentro da cadeia meromicólico. Esta variação na metilação e ciclopropanação dos ácidos micólicos pode ser importante para vários estilos de vida adotados por várias espécies de micobactérias. Proteínas responsáveis pela ciclopropanação sugerem redução na fluidez da parede celular, contribuindo para a impermeabilidade e proteção celular em ambiente hostil (Dubnau et al., 1997).

Em resposta a mudança de temperatura, por exemplo, as bactérias são capazes de modificar sua composição na membrana para manutenção da fluidez apropriada. O mecanismo mais comum usado pelos microrganismos para esta modulação é alterar a proporção da cadeia longa e ácidos graxos saturados dos glicerofosfolipídeos (Los e Murata, 2004).

A perturbação ambiental sofrida pela célula provoca alterações nas propriedades físicas dos lipídeos de membrana. Acredita-se que as células percebam estas alterações através de proteínas sensoriais presentes na membrana, que transferem o sinal para redes de transdução de sinal, com consequente regulação da expressão gênica. O tipo de lipídeos da membrana vai interferir diretamente na regulação da atividade da proteína de membrana, como a translocação de pequenas moléculas, canais de íon, proteína quinase associada ao receptor e proteínas sensoriais (Tokishita e Mizuno, 1994; Los e Murata, 2004).

Micobactérias patogênicas, como *Mycobacterium tuberculosis*, que apresentam crescimento lento, modificam seus ácidos micólicos por ciclopropanação (Yuan et al., 1995). Pouco se sabe, em nível molecular, sobre a modulação de sua composição e conteúdo em resposta às mudanças no ambiente de bactérias (Meniche et al., 2009). Em 1995, Russell et al sugeriram que a modificação nos lipídeos da membrana é importante na manutenção da integridade da membrana e funcionalidade diante das perturbações externas.

Em *E.coli*, a bactéria mostrou uma habilidade para se adaptar ao ambiente de baixo pH (pH=5,0 em 10^8 células mL⁻¹) e é provável que a modificação na composição do ácido graxo na membrana tenha exercido um importante papel na plasticidade fenotípica. A tolerância a acidez parece estar correlacionada ao conteúdo de ácido graxo ciclopropano na membrana e,

por isso acredita-se que o aumento do nível de ácido graxo ciclopropano possa aumentar a sobrevivência da célula exposta a baixo pH (Brown et al., 1997).

Curiosamente, em *C.pseudotuberculosis* 1002 somente na condição de acidez o gene *ufaA*, que codifica uma suposta sintase do ácido graxo ciclopropano, não foi considerado diferencialmente expresso. Nas outras condições foi significativo ($pvalue < 0,001$) e considerado induzido, mas a razão (*fold-change*) entre estresse e controle foi baixa, praticamente a mesma expressão entre eles (Tabela 6). Contudo, o ácido graxo ciclopropano são sintetizados preferencialmente quando a bactéria entra na fase estacionária do seu ciclo de crescimento (Cronan, 1968). Em outro trabalho, (Mcgarrity, 1981) foi incapaz de demonstrar a regulação do ácido graxo ciclopropano em células na fase exponencial. Possivelmente, em *C.pseudotuberculosis* 1002 tenha ocorrido influência da densidade celular na regulação da expressão gênica, pois em todas as situações a expressão foi próxima ao controle, sem grandes alterações.

O gene Cp1002_1953 codificante de cutinase (ortólogo ao gene Rv3802 em Micobactérias) (Tabela 7) foi considerado diferencialmente expresso ($p-value < 0,001$) e considerado induzido (*fold-change* > 1) em todas as condições testadas. Uma pesquisa por busca de homologia já mostrou que Rv3802 está presente em todos os membros da subordem Corynebacterineae, mas não em outras bactérias. Em *Mycobacterium*, o gene está localizado a *upstream* aos três genes, *pks13*, *accd4* e *fadD32*, cujos produtos são requeridos para o passo final de condensação da biossíntese do ácido micólico (Meniche et al., 2009).

Em *C.pseudotuberculosis* 1002, como mostrado na figura 16, os genes estão nesta mesma posição, comprovando que o gene Cp1002_1953 seria mesmo o ortólogo de Rv3802, que codifica uma suposta cutinase (na verdade, lipase) a qual participa do metabolismo do ácido micólico. Em *C. glutamicum*, o gene ortólogo NCgl2775 mostrou estar envolvido na regulação da razão entre ácido micólico e fosfolipídeos em resposta a exposição ao aumento de temperatura da célula (42°C). O aumento da razão entre corynemicolatos saturados e insaturados sugere que a modificação lipídica ajuda a bactéria a se proteger contra prejuízos por aquecimento. Levando a supor que este gene e seus ortólogos, exerçam um papel na regulação da composição lipídica da membrana externa de *Corynebacterineae* sob condição de estresse (Meniche et al., 2009).

A enzima Culp6, codificada pelo gene Rv3802, não é exportada totalmente da célula, mas é retida na parede celular em *M. tuberculosis*. A posição no cromossomo próximo a genes relacionados ao metabolismo de lipídeos sugere papel direto na modificação da parede ou

modificação do lipídeo. O estudo da atividade enzimática mostrou que a enzima anotada no genoma como cutinase, não apresentou esta atividade, mas sim de lipase (West et al., 2009).

Em *C.pseudotuberculosis* 1002, os outros genes envolvidos na biossíntese do ácido micólico, vizinhos da então enzima anotada como cutinase, são *pks*, *pccB* e *fadD*, que se mostraram diferencialmente expressos. Além destes, outros genes envolvidos em modificação lipídica foram observados (tabela 7) e ainda que o *fold-change* não tenha sido, em sua maioria, maior que dois, isso já seria indício da tentativa de escape dos estresses no início da fase exponencial, o que leva a crer que a resposta é imediata.

Existe um estudo que identificou o gene *pks13*, em Micobactérias, codificante de uma condensase, enzima envolvida na última etapa da reação de condensação de duas cadeias de ácido graxo, levando a formação do micolato completo (Portevin et al., 2004). Os substratos da condensase são ativados pela acil-CoA sintase, que em outro trabalho revelou ser uma acil-AMP ligase (Trivedi et al., 2004), codificada pelo gene *fadD32* e acil-CoA carboxilase, codificada pelo gene *accD4*. A enzima, inicialmente denominada Pks13, mostrou ser essencial para sobrevivência de micobactérias e possivelmente também para Corynebacterineae. Assim, pode ser considerada como alvo específico para o desenvolvimento de drogas para o tratamento de infecções causadas por estas bactérias patogênicas (Portevin et al., 2004).

Em *C.glutamicum*, uma linhagem em que foi deletado o gene *pks13* mostrou ser deficiente na produção de ácido micólico, mas foi capaz de produzir os precursores de ácidos graxos (Portevin et al., 2004). No caso da *C.pseudotuberculosis* 1002, os ortólogos *pks*, *pccB* e *fadD* provavelmente exercem estes papéis. O gene *fadD* no estresse osmótico e acidez, codificante da acil-AMP ligase que converte ácido graxo de cadeia longa para acil-CoA, não foi considerado significativo. No estresse por temperatura, o gene foi induzido em uma razão de 1,2342 vezes em relação ao controle (tabela 7). Ela ativa o ácido graxo de cadeia longa como acil-adenilato e transfere para a Pks para uma extensão ainda maior. Assim, além da Pks13, Fad32 e Accd4 mostraram ser promissores como alvos para desenvolvimento de novas drogas antimicrobianas contra espécies patogênicas de micobactérias e microrganismos relacionados (Portevin et al., 2005).

O gene Cp1002_1530, codificante de uma suposta *long-chain-fatty-acid-CoA ligase*, também se mostrou induzido entre os envolvidos no metabolismo de ácido graxo. Ela apresenta como nome alternativo Acyl-CoA synthetase. Atua tanto em ácido graxo saturado como insaturado, catalizando a reação de produção de adenosina monofosfato+ difosfato+ acil-CoA (<http://www.expasy.org/enzyme/6.2.1.3>). A função precisa desta enzima não foi ainda

estabelecida. Na condição de acidez e osmótico o gene se mostrou induzido (tabela 7). No macrófago, um dos principais mecanismos de defesa da bactéria, pela incapacidade da ação das células fagocíticas, é justamente a presença de camada lipídica que a torna resistente (Bogdan et al., 1997). Células adaptadas à acidez de *Streptococcus mutans* mostraram ser enriquecidas em mono-insaturados e longas cadeias de ácido graxo comparadas às células não adaptadas (Fozo e Quivey, 2004). Por estes resultados, confirmamos que *C.pseudotuberculosis* 1002 apresenta vários genes envolvidos no escape da resposta do hospedeiro, que são acionados no início da fase exponencial (DO=0,2).

A modificação lipídica da proteína bacteriana facilita o ancoramento de proteínas hidrofílicas em superfícies hidrofóbicas, através da interação do grupo acyl ligado ao fosfolípido da parede, o que permite ancoramento suficiente em ambiente aquoso (Kovacs-Simon et al., 2011). Lipoproteínas bacterianas vêm mostrando vários papéis importantes como, absorção de nutrientes, transdução de sinal, adesão, conjugação, transporte (sistema de transporte ABC) e dobramento de proteínas extracitoplasmática. Em bactérias patogênicas, lipoproteínas mostram papel direto em funções associadas a virulência, como colonização, invasão, escape da defesa do hospedeiro e imunomodulação (Alloing et al., 1994).

Lipoproteínas são inicialmente traduzidas como pré-prolipoproteínas, que possuem a característica típica de proteínas secretadas que é a presença do peptídeo sinal com N-terminal (Inouye et al., 1977). Uma sequência conservada na extremidade C terminal do peptídeo sinal (lipobox) sofre uma modificação que é catalizada pela enzima lipoproteína diacilgliceril transferase (Lgt), resultando em prolipoproteína, consistindo no motivo diacilglicerol ligado ao grupo tioéster ligado na proteína (Babu et al., 2006). Após a lipidação, a lipoproteína peptidase sinal (Lsp) cliva o peptídeo sinal da prolipoproteína deixando a cisteína do lipobox com um novo resíduo amino-terminal, formando a lipoproteína madura (Tokunaga et al., 1982). Para melhor compreensão observa-se a figura 26.

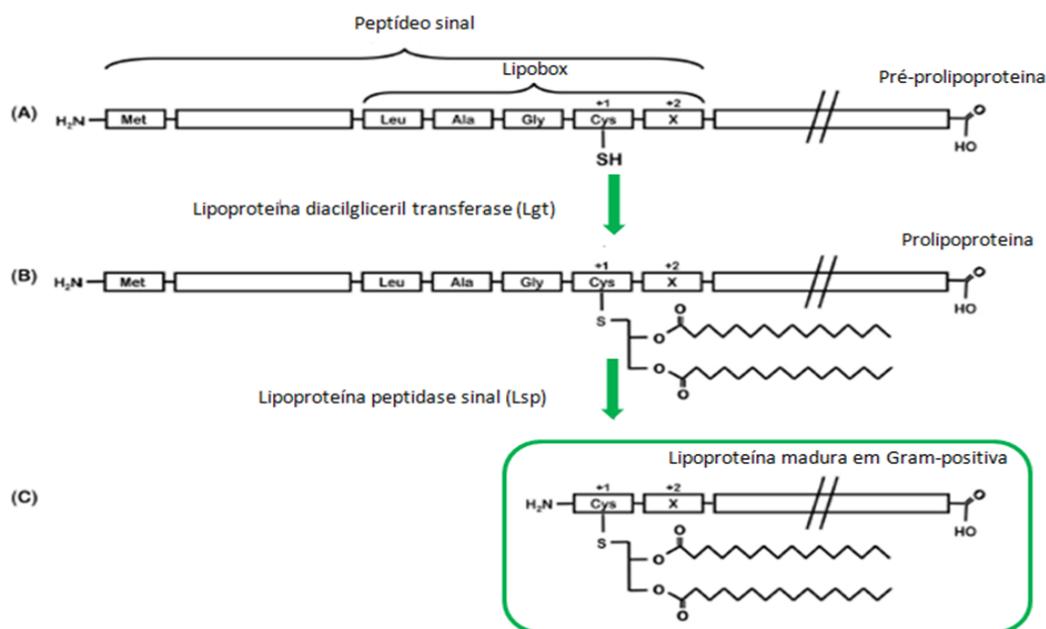


FIGURA 26 - Biossíntese da lipoproteína bacteriana.

(A)- O precursor de lipoproteína é a prelipoproteína, com peptídeo sinal na extremidade N-terminal que possui uma sequência consenso característica de lipobox. (B)- Durante a maturação da lipoproteína, o grupo tiol da cisteína do lipobox é modificado pelo motivo diacilglicerol através da lipoproteína diacilgliceril transferase (Lgt), que serve como ancorador de membrana. (C)- Depois da lipidação, a lipoproteína peptidase sinal (Lsp) cliva o peptídeo sinal, deixando a cisteína com novo resíduo amino-terminal, formando a lipoproteína madura em Gram-positiva. Figura adaptada (Kovacs-Simon, Titball, Michell, 2011).

Em Gram-positivas o precursor é translocado através do citoplasma direcionado pelo peptídeo sinal onde é modificado como descrito acima e a lipoproteína madura é então exposta ao ambiente extracelular. O atípico ácido micólico da parede de bactérias Gram-positivas, como em *Mycobacterium tuberculosis*, serve como potencial ponto de ancoramento para as lipoproteínas. Estas proteínas vêm demonstrando papel na virulência permitindo a interação com hospedeiro a partir da superfície de adesão e iniciação do processo inflamatório pela translocação de fatores de virulência dentro do citoplasma do hospedeiro (Das et al., 2009).

O papel da lipoproteína na virulência vem sendo observado através de estudos mutagênicos dos genes *lgt* e *lsp*, onde os mutantes apresentaram redução na adesão e internalização no hospedeiro, como nos mutantes para *lsp* em *Streptococcus pyogenes* onde se observou redução na ligação à fibronectina e laminina (Elsner et al., 2002). Em *Streptococcus pneumoniae*, o mutante para *lgt* se mostrou avirulento em modelos de camundongos (Petit et al., 2001), enquanto que o rompimento do *lgt* em *Streptococcus equi* resultou em moderada atenuação na virulência (Hamilton et al., 2006).

Em *C.pseudotuberculosis* 1002 observa-se que a condição de estresse influencia na expressão dos genes envolvidos na formação da lipoproteína. Ainda existe a influência,

provavelmente, da densidade celular, uma vez que, alguns genes não apresentaram uma expressão elevada. O gene *lgt* não foi considerado significativo na condição de acidez, mas apresentou-se induzido no estresse osmótico e térmico, em concordância com o resultado de *lppL* (suposta prolipoproteína), onde também na acidez não foi considerado diferencialmente expresso (tabela 7). O *lsp* que codifica a proteína que cliva o peptídeo sinal para então formar a lipoproteína madura foi reprimido no estresse osmótico e ácido, mas a indução no térmico apresentou um *fold-change* que demonstra uma expressão muito similar ao controle (tabela 7). O que pode ser indício que a lipoproteína madura nestas condições não ocorre com frequência nesta fase de crescimento.

Em Gram-negativa, *Salmonella typhimurium*, o gene *lgt*, juntamente com outros genes envolvidos na modificação lipídica de prolipoproteínas, demonstrou ser essencial para crescimento normal, divisão celular e viabilidade (Gan et al., 1993). Já em Gram-positiva, como *Bacillus subtilis*, não é necessário para o crescimento e nem para viabilidade, embora a lipoproteína PrsA seja essencial (Tjalsma et al., 1999) e em *S. pneumoniae*, a enzima codificada pelo gene *lgt* não é essencial para o crescimento *in vitro*, mas é essencial para viabilidade durante a infecção (Petit et al., 2001). O que nos infere que a resposta não é conservada entre os organismos e a regulação deve ser complexa.

Genes citados na literatura como prováveis candidatos a codificar lipoproteínas relacionadas à virulência, não foram considerados significativos em *C.pseudotuberculosis* 1002, como o gene codificante da proteína do sistema transportador do tipo ABC nomeada PstS, que é uma lipoproteína que contribui para virulência e tem demonstrado ser promissora como candidato vacinal (D'souza et al., 2002).

O gene Cp1002_0517 codificante da provável lipoproteína LpqB foi induzido em todas as condições, porém sua função ainda é desconhecida (Möker et al., 2004). Intrigante é que ele está localizado abaixo dos genes *mtrA2* e *mtrB* e nos organismos onde está presente, o stop códon do *mtrB* sobrepõe o start códon de *lpqB*, fortemente sugerindo que a tradução destes dois transcritos (*mtrB* e *lpqB*) ocorre ao mesmo tempo (Hoskisson e Hutchings, 2006). E realmente o gene *mtrB* foi induzido nas três condições e o *mtrA* não foi considerado significativo no estresse osmótico e térmico (não mostrados, pois o *fold-change* está abaixo do corte). Estes dois genes (*mtrA2* e *mtrB*) parecem codificar sistema de dois componentes e Hoskisson e Hutchings (2006) sugerem que a LpqB intervenha na transdução de sinal através da MtrAB pela interação com o domínio sensor MtrB, justamente por existir relatos de

lipoproteínas regular as quinases sensoras em outros organismos (Hutchings et al., 2000) Dartois et al., 1997), contudo são hipóteses.

Tem sido sugerido o papel de MtrAB na regulação de osmoprotetores, homeostase do envelope celular e progressão no ciclo celular. A impressão geral é que MtrAB percebe e responde ao estresse que pode interromper a progressão do ciclo celular e a MtrB (sensor) parece uma quinase bifuncional que reduz o nível de fosfo-MtrA na ausência do sinal (Fol et al., 2006). No entanto, como em *Mycobacterium tuberculosis* a MtrB se mostrou não essencial, acredita-se que MtrA (regulador) possa ser fosforilada de outra forma, na ausência de MtrB (Hoskisson e Hutchings, 2006) .

O gene *mtrA* em *Mycobacterium tuberculosis* revelou ser constitutivo durante o crescimento *in vitro* (Zahrt e Deretic, 2000) e dez vezes mais induzido em *M. bovis*, sob infecção em macrófagos, sugerindo que MtrA pode exercer um papel na proliferação em *Mycobacterium tuberculosis*. Em nosso estudo, a única condição em que o gene foi considerado significativo e induzido, em *C.pseudotuberculosis* 1002, foi no estresse ácido, onde simula-se uma condição do ambiente de macrófago, porém o *fold-change* foi um pouco acima de 1.

É de suma importância e essencial desvendar o papel destas supostas lipoproteínas, uma vez que em determinadas espécies demonstram importante papel na virulência e bons candidatos vacinais, porém estes resultados obtidos são indícios que nesta fase de crescimento, a atividade não parece relevante. Mesmo na condição controle, os genes não demonstram altamente induzidos, ou reprimidos. Possivelmente, são expressos em fase mais tardia de crescimento, mas alguns já demonstram uma importância com um valor de expressão bem acima do controle, e ainda demonstram ser regulados pelo ambiente e geralmente, genes com esta característica, são potenciais fatores de virulência (Mahan et al., 1993).

6.2. Core Estimulon entre os três estresses, com *fold-change* ≥ 2 ou $\leq 0,5$

Para entendermos o processo da infecção e revelarmos a modulação da resposta dentro da célula frente aos estresses, era preciso caracterizar os genes diferencialmente expressos e desvendar o papel funcional na célula. Assim, neste item, os genes que apresentaram variação na expressão (*fold-change*) alterada em no mínimo duas vezes (induzido ou reprimido) nas três condições ao mesmo tempo (core estimulon), foram caracterizados.

Entre os genes induzidos com *fold-change* ≥ 2 trinta fizeram parte do *core estimulon*. Pode-se observar na tabela 8 que o gene Cp1002_1418, mais expresso nas três condições, codifica uma proteína hipotética. O valor de RPKM deste gene no controle foi zero, ou seja, não obteve nenhuma expressão nesta condição, a que tudo indica. Esta proteína está anotada no genoma com um tamanho de 84 bases e 27 aminoácidos. No sequenciamento de RNA (e mesmo em outras tecnologias envolvidas na detecção de expressão gênica), considerando o tamanho dos transcritos, os maiores são preferencialmente mais detectados em relação aos mais curtos, pois a probabilidade de se ter mais representantes (fragmentos) na amostra é maior. Em outras palavras, o número total de leituras mapeando um determinado transcrito está diretamente ligado ao nível de expressão multiplicado pelo tamanho deste transcrito (Roberts et al., 2011).

Sendo assim, quanto maior o transcrito, mais leituras serão mapeadas, quando comparado ao transcrito menor que tenha apresentado uma expressão similar. Mesmo com um tamanho pequeno, este gene foi representado, indicando que o programa utilizado conseguiu superar o viés do tamanho e em questão biológica, esta proteína pode ser um alvo para estudos futuros. Entre os mais induzidos e reprimidos em todas as condições, estão as proteínas hipotéticas. Confirma-se assim, como este microrganismo é pouco caracterizado e justamente por isso, mais pesquisas serão necessárias para desvendar esta bactéria e alcançar o objetivo final de toda a pesquisa que seria o controle da doença, ou até mesmo erradicação.

Outro gene que apresentou uma expressão que chamou atenção foi o Cp1002_0730, codificante do regulador transcricional da família TerR, o qual foi mais induzido no estresse osmótico. Os membros da família são repressores e este gene está envolvido no controle transcricional de bomba de efluxo, rota de biossíntese de antibiótico, resposta ao estresse osmótico e moléculas tóxicas, controle da rota catabólica, processo de diferenciação, patogenicidade. A sobrevivência em ambiente instável depende de uma resposta rápida e adaptativa que se arma por proteínas regulatórias. Em muitas situações, as respostas são mediadas por reguladores transcricionais (Ramos et al., 2005). A maioria dos reguladores transcricionais envolvidos no controle transcricional são proteínas com dois domínios, um receptor de sinal e um domínio ligante ao DNA que transfere o sinal (Babu e Teichmann, 2003). Em outros reguladores, o receptor de sinal que arma o processo transcricional, envolve duas proteínas, como no sistema regulatório de dois componentes. Cerca de 95% de todos os fatores transcricionais descritos em procariotos apresentam motivo HTH (hélice-volta-hélice)

para se ligar ao alvo DNA (Argos et al., 1983), como apresenta o regulador TerR (Kendall et al., 2007).

Em *C.pseudotuberculosis* 1002, o gene se mostrou mais induzido no estresse osmótico (tabela 8), e neste caso é provável que esteja assumindo sua função envolvida na resposta ao estresse osmótico. Provavelmente controlando a saída de moléculas através da bomba de efluxo tentando atingir o equilíbrio para conseguir se manter no ambiente de alta salinidade. Ele se mostrou altamente induzido em todas as situações de estresse e este regulador está envolvido em muitas respostas, por isso, é de suma importância que estudos envolvendo a descoberta do seu regulon em *C.pseudotuberculosis* 1002 sejam realizados. Determinar a interação com outros genes, efeito na persistência da bactéria no ambiente, entre outras funções e, assim, abrir o leque de candidatos para controle da doença.

O gene Cp1002_0758, codificante de uma suposta Nudix hidrolase, também se mostrou induzido nos três ambientes, com valor de *fold-change* elevado nas três condições, principalmente no estresse osmótico (tabela 8). Nudix hidrolases, caracteriza-se como uma ampla família de enzimas que catalisam a hidrólise de uma variedade de nucleosídeos difosfato. Depende da preferência pelo substrato e a especificidade de substrato é utilizada para definir as superfamílias. Estes substratos são metabólitos ou moléculas de sinalização celular que requerem regulação durante as diferentes fases do ciclo celular, ou durante períodos de estresse. Em geral, o papel da Nudix hidrolase é de limpeza, ou manter a viabilidade das células, servindo, assim como vigilância e "casa de limpeza". Os substratos de Nudix hidrolase, são supostamente "alarmones" que são sintetizados em alta concentração na presença de estresse (Bochner et al., 1984).

Em *E.coli*, estes *alarmones* ligam-se em proteínas de choque térmico e oxidativo e acredita-se que modulam a resposta aos estresses (Fuge e Farr, 1993). *Alarmones* é uma molécula de sinalização intracelular que é sintetizada em alto nível quando as células enfrentam diferentes tipos de agressões ambientais. O típico *alarmones* é (p)ppGpp (guanosina pentafosfato ou tetrafosfato), cuja produção é desencadeada por diferentes sinais, incluindo ausência de aminoácido (Dalebroux et al., 2010). O acúmulo de (p)ppGpp resulta em repressão dos fatores envolvidos no crescimento celular e indução daqueles requeridos para adaptação ao estresse, incluindo fatores críticos para virulência (Rohmer et al., 2011).

Embora originalmente a Nudix hidrolase tenha sido descrita apenas como um mecanismo de prevenção de mutação, como MutT em *E.coli*, agora é evidente que apresenta funções diferentes em vários hospedeiros, incluindo degradação de compostos intracelulares

tóxicos e patogênese bacteriana (Bessman et al., 1996; Bessman et al., 2001). Em *Legionella pneumophila*, o gene nomeado *nudA*, codificou uma Nudix hidrolase e o mutante *nudA* cresceu mais lentamente no macrófago que a selvagem e apresentou tamanho menor da colônia. O gene mostrou ser de grande importância para resistir ao estresse em *L. pneumophila* e revelou ser um fator de virulência (Edelstein et al., 2005). Este gene deve ser analisado quanto ao seu potencial vacinal, pois demonstra ser de grande importância para permanência de *C.pseudotuberculosis* 1002 em condições alteradas do ambiente.

O gene Cp1002_1870 é codificante de uma suposta sortase associada a fimbria. Esta enzima está envolvida na ligação covalente de algumas proteínas de superfície da bactéria ao peptidoglicano, exercendo papel fundamental na apresentação de proteínas de superfície e na virulência do patógeno, como em *Staphylococcus aureus* (Cossart e Jonquières, 2000). As sortases são transpeptidases cisteína encontrada em bactérias Gram-positivas que ancora proteínas de superfície para o peptidoglicano da parede celular. Um critério inclui a presença de um motivo conservado (LPXTG) que contém resíduo de cisteína (Scott e Barnett, 2006). A maioria destas proteínas são ancoradas pela sortase “housekeeping”, designada SrtA. Assim, o substrato da proteína de superfície é clivado no sinal conservado da parede celular e covalentemente ligado ao peptidoglicano para ser apresentado na superfície (Scott e Barnett, 2006).

Existem várias classes e subfamílias para as sortases. Neste caso de indução em *C.pseudotuberculosis* 1002, a enzima apresentou homologia à sortase C, que desempenha papel na patogênese em *Streptococcus pneumoniae*, e é requerido para formação do pili na superfície de *C.diphtheriae* (Ton-That e Schneewind, 2003). A presença de pili na superfície das células bacterianas e seu papel na aderência tornam-os candidatos ideais para vacina e sua imunogenicidade já é bem documentada (Mora et al., 2005) (Barocchi et al., 2006).

Em *C.pseudotuberculosis* 1002, a enzima apresentou uma expressão maior no estresse osmótico, seguido do estresse térmico. A simulação da acidez no macrófago foi onde teve menor expressão, mas ainda o dobro em relação ao controle. O contato inicial da bactéria patogênica com células do epitélio no hospedeiro, por exemplo, ocorre por meio de pili (Webb e Kahler, 2008), onde, após invadir o hospedeiro, precisa aderir à célula do hospedeiro para então estabelecer a infecção. Assim, este gene torna-se importante alvo para testes vacinais.

O gene *dtl* codifica a suposta D-tyrosyl-tRNA(Tyr) deacylase. A inibição do crescimento celular por diversos D-aminoácidos podem ser explicada por produção *in vivo* de moléculas de D-aminoacil-tRNA. *E.coli* e células de levedura expressam genes codificantes da

enzima D-tyrosyl-tRNA(Tyr) deacylase, capaz de reciclar tais moléculas D-aminoacil-tRNA para tRNA livre e D-aminoácido (Ferri-Fioni et al., 2001). Tem sido demonstrado que a síntese de proteína mediada por ribossomos é limitada ao grupo L-aminoácido. Uma das descobertas biológicas mais surpreendentes nos últimos anos foi a detecção de D-aminoácidos em diversos organismos vivos, sendo moléculas mais ativas do que se imaginava (Ferri-Fioni et al., 2001).

Microrganismos unicelulares produzem, metabolizam e utilizam D-aminoácidos (Hopkins et al., 1997), como na parede celular que contribuem para a resistência a digestão por enzimas proteolíticas (Yang et al., 2003). A D-tyrosyl-tRNA(Tyr) deacylase é produzida para ser uma enzima editora, que remove o D-tyrosina e outros D-aminoácidos do tRNA carregado, evitando assim a incorporação incorreta do D-aminoácido dentro de proteínas (Yang et al., 2003). O envolvimento da enzima na proteção contra D-Tyr ou D-Trp, ou D-Asp-tRNA foi demonstrado pela acentuada toxicidade de cada um destes três D-aminoácidos em resposta a inativação do gene *dtd* (Soutourina et al., 2000). Em *C.pseudotuberculosis* 1002 seria interessante realizar estudos específicos para descobrir a verdadeira função desta enzima na bactéria, mas a que tudo indica promove proteção contra a toxicidade das moléculas D-aminoácidos que podem se formar na célula bacteriana.

O gene Cp1002_0025 codifica uma suposta CRISPR-associated protein. Repetição palindrômica curta clusterizadas interespaçadas regularmente (CRISPR), descreve uma classe de DNA repetitivo encontrada no genoma de bactéria e arquea. Não codificam proteínas, mas são transcritos e processados se tornando um small RNA que pode exercer diversas funções, (Haft et al., 2005). O cluster protege muitas bactérias e arqueas contra invasão de plasmídeo, transposase e vírus. O procarioto adquire resistência ao vírus, pela integração do fragmento curto do DNA viral dentro do cluster (CRISPR). A sequência integrada funciona como memória, evitando a infecção no hospedeiro (Haft et al., 2005).

Um número de genes associados (*cas*) ao CRISPR tem sido relatado como essencial para fenótipo resistente ao fago. A proteína associada (Cas) usa a sequência integrada para mediar a resposta antiviral. Após a transcrição do CRISPR, o complexo da proteína Cas, nomeado Cascata, cliva o RNA precursor de CRISPR em cada repetição e mantém o produto da clivagem contendo a sequência derivada do vírus, formando assim o RNA maduro crRNA e tudo indica que a formação do RNA maduro pela cascata seja o mecanismo requerido para a defesa evitando a predação por fago. No entanto, o mecanismo molecular desta adaptação e sistema de defesa em procariotos ainda é desconhecido (Brouns et al., 2008).

Algumas proteínas Cas mostram similaridade com helicases e proteínas de reparo, embora a função da maioria ainda seja desconhecida. Esta família pode ser subdividida de acordo com a organização do operon e filogenia (Haft et al., 2005). No genoma de *C.pseudotuberculosis* 1002 somente um representante da família, Cas5, está presente e está localizado dentro da ilha de patogenicidade. O gene da Pld está localizado *downstream* em uma distância de 507 aminoácidos e na fita oposta, uma transposase está há uma distância de 24 aminoácidos do gene. Mais detalhes funcionais em *C.pseudotuberculosis* 1002 serão necessários para conhecer a importância desta proteína na resistência aos estresses.

Entre os mais induzidos nas três situações chamaram atenção os genes Cp1002_0457 e Cp1002_0729, que codificam major facilitator superfamily permease (MFS). É um grande e diverso grupo de transportadores secundários que incluem uniporter, symporter e antiporter. Estes transportam pequenos solutos em resposta ao gradiente de íons quimiosmótico (<http://pfam.sanger.ac.uk/family/PF07690>). As proteínas desta família facilitam o transporte através do citoplasma ou membrana interna de uma variedade de substrato incluindo íons, açúcar fosfato, nucleosídeos, aminoácidos, peptídeos, drogas, facilitando a passagem que, por difusão simples, demorariam muito tempo para atravessar a membrana de modo a igualarem suas concentrações (Pao et al., 1998).

No caso do estresse osmótico, em *C.pseudotuberculosis* 1002, possivelmente foi induzido, tentando estabelecer um equilíbrio entre moléculas para evitar prejuízo da célula e foi exatamente neste ambiente que o *fold-change* se apresentou com maior valor, com quase 5x a mais de expressão em relação ao controle (tabela 8). Inclusive em eucariotos (planta), recentemente confirmaram o papel dos carboidratos na tolerância ao sal, lutando contra o estresse osmótico. Na análise do transcriptoma e metaboloma encontraram forte acúmulo de açúcar primário, ativação de rota para produção de açúcar e rápido consumo de metabólitos secundários (Janz et al., 2010). No estresse por acidez, a permease provavelmente esteja atuando para equilibrar os íons e manter um ambiente menos acidófilo, para tolerar esta condição. No estresse por choque ácido e térmico, o *fold-change* foi praticamente o mesmo.

A indução da resposta ao estresse térmico, melhora a termotolerância, tolerância ao sal e até mesmo aos metais pesados (Inbar e Ron, 1993), pois a resposta ao choque térmico não é limitada às mudanças na temperatura, mas é uma resposta ao estresse geral (Gophna e Ron, 2003). Sendo assim, é possível acreditar que uma resposta ao choque térmico em *C.pseudotuberculosis* 1002, traz benefícios também diante do estresse salino, por exemplo, uma vez que a bactéria se depara com estas condições ao mesmo tempo no hospedeiro. Não

tem relatos da função exata desta permease no estresse térmico, mas seria interessante estudar mais profundamente este gene, pois pode vir a ser um alvo importante no desenvolvimento de uma vacina, uma vez que está presente nos três estresses em uma indução com *fold-change* maior que o dobro em relação ao controle.

O gene *phuC* codifica uma suposta Iron(III) *dicitrate transport permease-like protein* yusV, cujo domínio está relacionado ao sistema de transporte de ferro do tipo ABC, envolvido na absorção de sideróforo, grupo heme, vitamina B12. A aquisição do ferro é, provavelmente, o principal determinante para descobrir se um microrganismo é capaz de se manter dentro do hospedeiro. Sem esta habilidade, o microrganismo será incapaz de crescer e será eliminado pelo sistema de defesa do hospedeiro, ou morrerá por falta de nutrientes. A aquisição do ferro é reconhecida como um dos passos fundamentais para o desenvolvimento de qualquer patógeno no hospedeiro (Ratledge e Dover, 2000).

Em *C.pseudotuberculosis* 1002, este gene parece contribuir imensamente a sua permanência diante das perturbações enfrentadas que do contrário, seria prejudicial. O gene mostrou um valor quatro vezes a mais que o controle, e os valores estão bem próximos nas três condições (tabela 8). O gene está localizado dentro de uma das ilhas de patogenicidade, no genoma de *C.pseudotuberculosis* 1002. Em um ambiente onde o ferro não está totalmente disponível, como no hospedeiro, este gene deve ser um dos mais expressos para tentar solucionar o problema.

O gene *corA* codificante da suposta proteína magnesium/cobalt transporter, também está entre os genes comuns entre os três estresses, com *fold-change* acima de dois. O magnésio é o mais abundante cátion divalente dentro de procariotos e eucariotos e apresentam muitas funções na célula. Contribui com a estabilidade da membrana e é um cofator com moléculas como ATP (Jou e Cowan, 1991), tRNA (Quigley et al., 1978) e enzimas como ribonucleaseH (Jou e Cowan, 1991). O prejuízo ao RNA pode estar associado com queda de pH, e pode ser devido a remoção do magnésio, que é também responsável pela integridade ribossomal (Wesche et al., 2009). É considerado como molécula sinalizadora para o sistema sensor/sinalizador PhoP/PhoQ em *Salmonella* patogênica (Groisman et al., 1989). A bactéria precisa de concentrações micromolares de Mg^{2+} para o crescimento e tem necessidade de adquirir esta molécula para melhor adaptação no ambiente.

A classe CorA é a mais abundante classe dos transportadores de Mg^{2+} em procarioto. CorA é diferente das outras classes de transportadores de Mg^{2+} porque também transporta cobalto e níquel e pode mediar tanto o influxo quanto o efluxo destes íons (Van Vliet et al.,

2001). Apesar da importância, pouco se sabe sobre a regulação do magnésio e o mecanismo de transporte (Szegedy e Maguire, 1999). Em *C.pseudotuberculosis* 1002, o papel exato deste gene na célula deve ser estudado, pela importância que parece apresentar e analisar a possibilidade de se tornar um novo candidato a estudos vacinais.

Outros genes compartilhados entre os estresses foram os relacionados aos ribossomais. Para manter a replicação, ainda que desacelerada em relação ao controle, faz-se necessário manter os ribossomais que são indispensáveis ao crescimento, acrescido dos fatores de transcrição e chaperonas necessárias para assegurar adequada tradução e dobramento da proteína (Karlin e Mrázek, 2000). Genes que codificam proteínas ribossomais L14, L5, L24 foram diferencialmente expressos no *core estimulon*, indicando uma resposta relacionada ao crescimento indispensável em todos os estresses. Estas proteínas ligam-se na subunidade 50S (constituída pela subunidade 23S e 5S) contribuindo para a formação completa do ribossomo 70S, ao juntar-se com a subunidade 30S. A habilidade em se dividir rapidamente, possibilita a adaptação às mudanças no ambiente. A proteína L14 se liga no rRNA 23S e a L5 é uma das proteínas que permite a ligação do rRNA 5S na subunidade maior e tem importante papel na conformação do rRNA 5S (Karlin e Mrázek, 2000). Esta indução comprova que o crescimento não foi interrompido por completo, apenas reduzido, que foi confirmado pelo experimento de Unidade Formadora de Colônia (figura 5), uma estratégia para manter a viabilidade e garantir melhor adaptação a uma nova condição (Jozefczuk et al., 2010).

Vale destacar que os genes citados acima, em comuns nas três condições, não apresentaram um valor de RPKM muito expressivo, mas em relação ao controle apresentou um *fold-change* maior que duas vezes, justamente pelo fato do controle ter apresentado um RPKM muito baixo. Os transcritos estão sujeitos a degradação, que contribui para um ruído nos dados de cobertura. Notavelmente o nível de expressão é específico à condição de crescimento da célula. Em *Bacillus athracis* a média de cobertura mostrou que em torno de 15%-18% dos 5.661 genes representados, apresentaram uma distribuição de cobertura menor que dois, dependendo da condição de crescimento (Martin et al., 2010). Porém, ainda que em baixa cobertura, se visualizamos que os genes estão sendo induzidos e variando entre condições tão precárias para a bactéria, é bem provável que estejam, em sua maioria, envolvidos na virulência, pois a expressão coordenada de genes de virulência em muitas bactérias patogênicas é regulada por estímulos ambientais (Miller et al., 1989).

O gene *hemK* em *E.coli* foi sugerido inicialmente codificar *protoporphyrinogen oxidase* (Nakayashiki et al., 1995), porém uma análise funcional do homólogo em

Saccharomyces cerevisiae sugeriram que o hemK na verdade tem homologia com *S-adenosyl-methionine-dependent methyltransferase* ou *N⁵-glutamine methyltransferase* (Le Guen et al., 1999). O gene *hemK* está imediatamente downstream e co-expresso ao gene *prfA*. A deleção de *hemk* em *E.coli* k12 levou a um crescimento reduzido em meio rico e impediu a metilação do *peptide chain release factor 1* (RF1) (Heurgué-Hamard et al., 2002). Este gene está entre os transcritos que são necessários e importantes para o crescimento, pelo menos em *E.coli*, e em outros organismos resta ser elucidado. Os RFs apresentam um motivo conservado GGQ, e neste motivo, o resíduo glutamina é modificado pós-traducionalmente para *N⁵-methylglutamine* em ambos fatores RF1 e RF2. O gene *hemK*, mostrou a atividade de metilar a região no experimento *in vitro*. (Heurgué-Hamard et al., 2002).

Em *C.pseudotuberculosis* 1002, provavelmente o gene foi induzido com o objetivo de garantir a sobrevivência em ambientes tão precários, contribuindo para seu crescimento, mantendo a metilação dos fatores de liberação da cadeia (RF). Na realidade houve uma redução no crescimento das colônias, mas observamos a resistência da bactéria em condições tão desafiadoras.

O gene *prfA* codifica o suposto *Peptide chain release factor 1* (RF1). O término da tradução ocorre quando a elongação no ribossomo se depara com sinal de parada no mRNA. Duas classes de fator de liberação da cadeia do peptídeo (RF1 e RF2) são responsáveis pelo reconhecimento, sendo que RF1 reconhece a sequência UAA e UAG e RF2 reconhece UAA e UGA (Kisselev e Buckingham, 2000). Em *C.pseudotuberculosis* 1002, este gene, diferente dos outros relatados acima, apresentou alto valor de RPKM em todas as situações, e apresentou *fold-change* maior que dois. Interessante que apresentou valores muito próximos nas três condições, indicando a grande importância para a manutenção da bactéria. Os RFs apresentam sempre um único motivo, que é conservado entre todos os reinos Gly- Gly –Gln (GGQ). As duas glicinas são essenciais para a atividade e o resíduo glutamina pode ser substituído, com restrição parcial da atividade *in vitro*. O papel preciso desta região resta ser elucidado (Heurgue-Hamard et al., 2002).

O gene *ruvC* codificante da suposta *Crossover junction endodeoxyribonuclease*, apresentou também nos estresses, um valor de RPKM mais expressivo. A recombinação genética é uma poderosa força na evolução dos microrganismos patogênicos e que facilita a aquisição de resistência aos antibióticos, determinantes antigênicos e fatores de virulência, que resultam na patogenicidade alterada (Maiden, 1998). O processo de recombinação homóloga envolve troca da fita simples entre parceiros homólogos, levando a formação do chamado

junção de Holliday. A endonuclease RuvC (resolvase) reconhece a junção Holliday, resolve a junção por incisão da fita dupla em todo ponto da ramificação para liberar o duplex que pode ser selado pela DNA ligase (Dunderdale et al., 1991).

Foi construído uma *H.pylori* mutante para *ruvC*, e o resultado revelou que a inativação do gene conferiu redução na sobrevivência dentro do macrófago e que a proteína é essencial na persistência da infecção por *H.pylori*. O mutante apresentou-se mais sensível ao estresse oxidativo e morte intracelular no macrófago. Observou-se que a manutenção da integridade do DNA é vital para viabilidade nas células fagocíticas (Loughlin et al., 2003). Em *E.coli*, a inativação de *ruvA*, *B* ou *C* resultou em modesta redução na recombinação genética, mas um defeito significativo no reparo do DNA (Lloyd, 1991). Em *H.pylori*, a inativação do *ruvC* resultou em defeito no reparo de DNA similar aos resultados em *E.coli*, mas uma grande redução na frequência de recombinação.

A habilidade de reparo no DNA prejudicado pelo estresse oxidativo produzido dentro dos macrófagos, é essencial para uma alta virulência e sobrevivência de *Salmonella typhimurium* dentro dos macrófagos (Buchmeier et al., 1993). Pela facilidade de aquisição de novos genes a partir de fontes diversas e o rearranjo de genes existentes, a recombinação homóloga pode ser crítica para adaptação de antígeno de superfície em resposta as mudanças ambientais presentes no estômago. Especula-se que uma recombinação possa contribuir para a infecção crônica pela substituição de versões alternativas de adesinas em superfície, antígenos, promovendo a evasão imune (Loughlin et al., 2003).

Neste trabalho, este gene, a princípio, demonstra ser muito importante para a manutenção da bactéria em meios prejudiciais, além dos valores expressivos em todas as condições, o que nos leva a supor, que a bactéria tem recursos para sanar o prejuízo causado pelas alterações ambientais já no início da sua replicação. A condição de acidez apresentou um *fold-change* um pouco maior que nas outras condições (tabela 8), em concordância com o ambiente tendencioso a promover estresse oxidativo.

O último gene que se integrou no *core estimulon*, entre as três condições testadas é o *infA*, o qual supostamente codifica o Translation initiation factor IF-1. A função exata desta proteína ainda é desconhecida, mas já demonstrou ser essencial para viabilidade celular em *E.coli*. (Cummings e Hershey, 1994). Quando o nível de IF-1 reduz, forma-se menos polissomos e diminui o crescimento celular, o que fortemente sugere que IF-1 seja responsável pelo crescimento celular e tenha importante participação na fase inicial da tradução (Dahlquist e Puglisi, 2000).

Em *C.pseudotuberculosis* 1002 o gene apresentou quase 2000 de RPKM no controle e nas condições de estresses valores acima de 3000 RPKM. O que nos revela ser um gene que é requerido a todo instante, mas diante de uma situação de estresse, e de acordo com a necessidade da bactéria, a expressão é alterada para, possivelmente, tentar compensar o prejuízo causado pelo ambiente e, assim, manter a viabilidade celular.

Muitas proteínas hipotéticas, de membrana ou sem domínio anotado foram destaques entre os três estresses. Pode-se perceber quão pouco se sabe deste microrganismo e seu estilo de vida e muitos estudos serão necessários para desvendar os mistérios desta bactéria e a resistência ao sistema de defesa do hospedeiro no processo de infecção. Alguns genes não foram caracterizados por escassez de informação.

Entre os genes que foram reprimidos, apresentando um *fold-change* pelo menos duas vezes menor estão listados na tabela 9. Apenas dez genes foram selecionados neste corte.

O gene *argJ* codifica a suposta proteína bifuncional ornithine acetyltransferase/N-acetylglutamate synthase, que catalisa duas atividades envolvendo a biossíntese de arginina: síntese do acetilglutamato a partir do glutamato e acetil-CoA e a produção de ornitina pela transacetilação entre acetilornitina e glutamato. Catalisa o primeiro e o quinto passo para biossíntese da arginina, acetilação do glutamato, com deacetilação de N-acetilornitina, que permite reciclagem do grupo acetil na rota de biossíntese de arginina (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=73147>).

Em *C.pseudotuberculosis* 1002, a repressão deste gene possivelmente prejudicou a rota de biossíntese da arginina, que demonstra ser de suma importância em situações prejudiciais para a célula, como ser precursora de compostos orgânicos que podem reduzir as consequências causadas pelo estresse oxidativo (Cui et al., 2001). Arginina é sintetizada por rotas complexas que são reguladas pelo repressor ArgR (Tian G e Wk., 1994) e em *C.pseudotuberculosis* 1002 o gene *argR* foi reprimido, mas não foi considerado significativo em nenhuma condição, o que levaria a supor que a síntese de arginina não seria reprimida. Porém, a regulação deve ser dependente de outros fatores, como em *E.coli* que demonstrou ser regulada positivamente pela falta de nutrientes da fase estacionária e pelo RpoS que é um sigma alternativo de fase estacionária (Weerasinghe et al., 2006), sendo indício de que o operon de arginina apresenta uma expressão dependente da densidade celular.

A biossíntese de arginina é elevada em resposta a falta de nutrientes, estresse ou em ambientes onde a arginina é limitada e principalmente na fase estacionária, o nutriente é

bastante limitado e a arginina sendo um reservatório de nitrogênio, contribui extremamente na manutenção do organismo nesta fase (Weerasinghe et al., 2006).

Este gene pertence ao provável operon *arg* o qual fazem parte 8 genes, no genoma de *C.pseudotuberculosis* 1002. Os genes apresentaram variações na expressão em todas as condições, mas estão fora do corte estabelecido neste trabalho, o que sugere que sua maior expressão ocorra realmente em fase tardia.

O gene *nanK*, codifica o suposto produto N-acetylglucosamine quinase (GlcNAc quinase), o qual participa do processo metabólico, e pertence a classe de proteínas (enzimas) que catalizam a fosforilação do açúcar ou carboidrato (Weihofen et al., 2006). A GlcNAc deve ser fosforilada para entrar na rota metabólica de GlcNAc-6-fosfato, um componente utilizado na biossíntese de UDP-GlcNAc (UDP-N-acetylglucosamine) que será empregado para produzir glicosaminoglicano, proteoglicano e glicolípides ou entra no metabolismo de energia, e a GlcNAc quinase quem converte a GlcNAc para GlcNAc-6-fosfato (Uehara e Park, 2004).

A GlcNAc (*N*-acetylglucosamine) é o principal componente da parede celular e em *E.coli*, durante o crescimento, uma parte da parede é degradada e o composto GlcNAc-anhydro-*N*-acetylmuramyl é importado para o citoplasma e clivado para liberar GlcNAc e outros elementos. A GlcNAc ou é reciclada envolvendo a fosforilação, uma vez que GlcNAc-6-fosfato é utilizada para síntese de mureína (peptideoglicano), ou pode ser metabolizada pela glicólise (Uehara e Park, 2004).

Este gene reprimido é indício que a fosforilação não ocorreu neste início da fase exponencial e assim, houve queda no metabolismo de energia ou na biossíntese do peptideoglicano. Os substratos desta enzima são ATP e *N*-acetylglucosamine, enquanto que seus produtos são ADP e GlcNAc-6-fosfato (Weihofen et al., 2006).

A transferência do fosfato ocorre a partir do ATP e a síntese de ATP está em baixa também em *C.pseudotuberculosis* 1002, uma vez que os genes responsáveis pela sua síntese (ATP sintase) foram reprimidos, em sua maioria e não foram considerados significativos nas três condições, colaborando com a justificativa da queda na expressão deste gene. Obviamente outros fatores atuam na regulação deste gene e é importante que sejam revelados. Não foi encontrado na literatura o comportamento desta enzima diante de estresses para comparação com os resultados do nosso trabalho.

Um dos sistemas de transporte mais versáteis é o sistema Opp, pois transporta uma variedade de peptídeos. Cinco produtos gênicos são necessários, OppABCDF, para o funcionamento deste transportador. Dois deles, OppB (em Cp1002-ABC-type

dipeptide/oligopeptide/nickel transport system, secreted component) e OppC, são altamente hidrofóbico, proteína integral de membrana e responsáveis por mediar a passagem de peptídeos atravessando a membrana citoplasmática (Pearce et al., 1992). O OppD e OppF se ligam na parede interna na membrana citoplasmática, regulando hidrólise de ATP e geração de energia requerida para transportar o soluto (Monnet, 2003).

A proteína específica ligante OppA (em Cp1002-ABC-type dipeptide/oligopeptide/nickel transport system, permease oligopeptídeo) é responsável por reconhecer, ligar os substratos de peptídeos para o complexo permease associado a membrana, conferindo assim, especificidade e afinidade para o sistema de transporte (Higgins, 2001). Em estudo com *Streptococcus agalactiae*, a deleção do gene *oppB* reduziu sua aderência nas células epiteliais, reduziu expressão do gene *fb*s que codifica a proteína ligante ao fibrinogênio e por isso, é provável que este gene esteja envolvido na modulação dos traços de virulência e expressão dos genes de virulência nesta espécie (Nepomuceno et al., 2007).

O gene *oppA*, está envolvido com transporte de oligopeptídeo através da energia oriunda da hidrólise de ATP. Um estudo transcricional do genoma de *Yersinia pestis* com microarranjo de DNA, sob condição de alta osmolaridade (sorbitol) e alta salinidade (NaCl) na fase exponencial ($A_{260}=0,6$), demonstrou que o gene *oppA* foi considerado induzido apenas no meio sob alta osmolaridade, em alta salinidade apresentou expressão similar ao controle (Han et al., 2005).

Em *L.monocytogenes*, foi expresso predominantemente no meio da fase exponencial e mostrou-se requerido para o crescimento da bactéria em baixa temperatura e favoreceu sobrevivência intracelular em macrófago (Borezee et al., 2000). Trabalhando com linhagem selvagem e mutante *oppA*, estes autores sugeriram que OppA favoreça escape inicial do fagossomo e multiplicação intracitoplasmática dentro do macrófago. O que leva a crer que a absorção do peptídeo possa exercer um papel importante na sobrevivência intracelular da bactéria. Uma especulação é que o peptídeo acumulado através da permease possa proteger a bactéria no fagossomo, ou ainda, que os peptídeos absorvidos poderiam ativar uma rota de transdução de sinal desconhecida, modulando a cinética de expressão de genes de virulência requeridos para escapar dos fagossomos.

A limitação de nutrientes pode prejudicar a bactéria nestes ambientes alterados. No entanto, outras permeases de peptídeo presente na bactéria, podem compensar a absorção dos peptídeos. Em *C.pseudotuberculosis* 1002 tanto *oppA* quanto *oppB* foram reprimidos. Este gene deve ser regulado também por densidade celular, além de outros fatores desconhecidos.

O sistema Opp de *Salmonella typhimurium* e *Escherichia coli* estão envolvidos na absorção de nutrientes na reciclagem de peptídeos da parede celular para síntese de um novo peptídeo (Goodell e Higgins, 1987). Em *Vibrio fluvialis* o mutante *oppA* influenciou na absorção de peptídeo e afetou a produtividade de biofilme em meio BHI, sendo que a produção foi duas vezes maior que na linhagem selvagem, sugerindo papel negativo do sistema Opp na formação do biofilme (Lee et al., 2004). O sistema Opp mutante de *Streptococcus mutans*, e particularmente a proteína OppA, crescidas em meio TSBY (caldo tripticase soja suplementado com extrato de levedura) ou BHI, não representou um importante traço requerido para o crescimento e colonização, apresentando resultado contrário às outras espécies de *Streptococcus* (Nepomuceno et al., 2007).

Em *C.pseudotuberculosis* 1002 parece seguir a linha do mutante de *Streptococcus*, no sentido do gene reprimido não prejudicar o crescimento da bactéria. Outro sistema pode também estar envolvido para obtenção de peptídeos ideais para o desenvolvimento da bactéria nesta fase de crescimento, porém a redução no crescimento ocorre essencialmente em todas as situações de estresses como uma importante estratégia de ajuste da fisiologia celular para uma nova condição (Jozefczuk et al., 2010). Ainda pode-se sugerir um papel como em *Vibrio fluvialis*, sendo reprimido para favorecer uma formação de biofilme. O que parece é que o gene se comporta diferentemente a cada espécie e outros estudos específicos para *C.pseudotuberculosis* 1002 devem ser realizados, pois o sistema Opp já demonstrou, em outras espécies, seu importante papel na virulência e sobrevivência da bactéria em ambiente hostil.

A suposta proteína regulatória da família GntR, codificada pelo gene Cp1002_2063, apresenta como característica, um domínio N-terminal conservado envolvido com a ligação ao DNA e um domínio C-terminal variado, envolvidos na ligação efetora e/ou oligomerização (Vindal, Suma Ranjan, 2007). Em consequência desta variação no C-terminal os reguladores são classificados dentro de seis subfamílias (FadR, HutC, MocR, YtrA, AraR and PlmA) (Rigali et al., 2004).

Em *Bacillus subtilis* é conhecido como repressor do operon gluconato (Haydon e Guest, 1991). Em *C.pseudotuberculosis* 1002, o domínio anotado para esta proteína faz parte da família nomeada FCD, que é membro da subfamília FadR. Provavelmente se liga a uma série de moléculas efectoras que regulam a transcrição de genes através da ação do N-terminal ligante ao DNA. Estes domínios são reguladores do operons da biossíntese de açúcar. Na estrutura conhecida de FadR liga-se também a molécula acil-CoA (<http://pfam.sanger.ac.uk/family?acc=PF07729>). O regulador FadR possivelmente está

envolvido na regulação da expressão gênica em resposta a substratos oxidados, relacionadas ou ao metabolismo de aminoácido ou nas rotas metabólicas como glicolato, piruvato, lactato, malonato, gluconato (Vindal et al., 2007).

Agora, interessante é que os genes *nanK*, *oppA*, Cp1002_2063 (GntR) e *oppB*, envolvidos de alguma maneira ao metabolismo de energia, estão anotados próximos no genoma. Onde *oppA* e *oppB* estão na fita direta, a uma distância de 404 bases e *nanK* e Cp1002_2063 (GntR) estão na fita complementar em sobreposição, sendo que Cp1002_2063 está a uma distância de 372 bases de *oppA*, e todos foram reprimidos nas três condições. Pode ser que exista alguma relação entre eles, mas uma dependência da população bacteriana é provável que exista entre os fatores de regulação. Certos tratamentos, particularmente o térmico, levam a inativação de enzimas e rompimento do transporte ativo de cátions, açúcar, e aminoácido (Hurst, 1977), possivelmente estratégia de sobrevivência da bactéria.

6.3 Estimulon no estresse osmótico

Dentro dos estimulons será discutido os processos biológicos obtidos pelo Blast2GO por condição, porém alguns genes chamaram a atenção e serão discutidos individualmente entre induzidos e reprimidos.

6.3.1. Genes induzidos

Entre os genes induzidos, diferencialmente expressos, com o valor de corte pelo menos duas vezes mais expresso em relação ao controle, no meio suplementado com cloreto de sódio, detectamos muitas proteínas hipotéticas, inclusive sendo o gene mais induzido. Alguns genes nos chamaram a atenção não pelo valor de *fold-change*, mas pelo valor de RPKM, como o gene *spaD*, *vap*, Cp1002_1495 (entre outras hipotéticas) (tabela 10).

O gene *spaD*, codificante de uma suposta subunidade fimbrial, apresentou um *fold-change* quase 4 vezes maior em relação ao controle e é de extrema importância, pois já foi testado quanto a sua capacidade protetora em camundongos, se tornando possível candidato vacinal. Este gene foi interrompido em *C. pseudotuberculosis* linhagem virulenta T1 e no desafio desta linhagem mutante em camundongos BALB/c, conferiu 60% de proteção, resultado este obtido pela vacina comercial GlanvacTM3 (Dorella, Tese 2009). Esta proteína apresenta um domínio presente em *Staphylococcus aureus*, “Collagen-binding protein B

domain”, relacionado à aderência da bactéria (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=196798>).

A interação da bactéria com a célula do hospedeiro permite estabelecer uma relação patogênica. Esta adesão envolve organelas de superfície, incluindo pili ou fimbrias (Soto e Hultgren, 1999). O pili de bactéria Gram positiva é pouco estudado. A montagem destas estruturas em Gram-positivas requer presença da sortase, uma transpeptidase que cliva o precursor da proteína no motivo LPXTG e liga o C-terminal da proteína de superfície através de ligação amida, ao peptidoglicano (Ton-That e Schneewind, 2004). Em *C.pseudotuberculosis* 1002, o gene localizado abaixo do gene *spaD*, o Cp1002_1870, codifica uma sortase associada a fimbria, e na condição de estresse osmótico se apresentou 8 vezes mais expresso que o controle (tabela 8). Entender os detalhes moleculares da montagem do pili aumenta o entendimento da patogênese, estrutura e função de superfície microbiana, mas também fornece oportunidade terapêutica para inibição da função dos genes de virulência que levam ao estabelecimento da doença.

Este gene em *C.pseudotuberculosis* 1002 foi considerado diferencialmente expresso, com o *fold-change* ≥ 2 somente na condição de estresse osmótico, com alto valor de RPKM em relação ao controle. No estresse térmico não foi considerado diferencialmente expresso e na acidez foi induzido com valor de RPKM muito próximo ao controle, o que resultou em um *fold change* de 1,3 vezes.

O gene *vap*, que codifica uma suposta proteína associada a virulência, apresentou um RPKM no estresse osmótico acima de 1000 e *fold-change* de quase 4 vezes (tabela 10) a mais em relação ao controle e nas outras condições apresentou um *fold-change* próximo a 1 (anexo 8). Não tem descrição sobre este gene até o momento e segundo a pesquisa de domínio conservado do NCBI, consta de uma proteína ligante ao DNA pertencente a família de elementos de resposta xenobiótica do regulador transcricional (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=c109100>). Compostos tóxicos, não utilizados no metabolismo que se acumulam na célula são conhecidos como xenobiótico. O termo é também aplicado a substâncias presentes em concentrações muito mais elevadas que o nível normal. Eles são eliminados por um conjunto de enzimas específicas (Testa e Krämer, 2006). Tudo indica que o gene *vap*, em *C.pseudotuberculosis* 1002, participa da eliminação de compostos tóxicos na célula, ou até mesmo favorecendo para que sejam utilizados como fonte de energia, pois quando o xenobiótico se assemelha a estrutura de um composto natural, pode

ser reconhecido pelo sitio ativo da enzima, permitindo ser quimicamente transformado, tornando uma possibilidade nutritiva para o microrganismo (Alexander, 1999).

6.3.2. Genes reprimidos

Entre os reprimidos muitos genes codificantes de proteínas hipotéticas estão relatados. Novamente, descrevemos genes que apresentaram um alto valor de RPKM, porém agora no controle. Vários genes apresentaram esta característica (tabela 11), mas relatamos três que nos chamaram mais atenção pelo valor de RPKM em relação ao estresse.

O gene Cp1002_0519, codifica uma suposta proteína Y (pY) associada ao ribossomo cuja função está relacionada a resposta ao estresse, ligando na subunidade ribossomal e impedindo a tradução por não permitir a ligação do tRNA aminoacil no sitio A do ribossomo. Em *E.coli*, a proteína não foi detectada em condições fisiológicas normais, porém foi induzido em temperaturas baixas. Na fase estacionária, fase de suspensão do crescimento, o ribossomo apresentou a mesma quantidade de proteína Y detectada nas células sujeitas ao estresse de queda de temperatura por uma hora, indicando ser induzida por estresse e também por alta densidade celular (Agafonov et al., 2001).

Em *C. pseudotuberculosis* 1002, o gene foi reprimido no estresse osmótico quase 5 vezes em relação ao controle, onde se apresentou altamente expressa. Este processo nos leva a concluir que na condição de estresse a bactéria *C. pseudotuberculosis* 1002 foi menos prejudicada quanto a interrupção de síntese de proteína que no controle. Em *E.coli*, quando o crescimento da bactéria é retomado, a proteína não é mais detectada no ribossomo (Agafonov et al., 2001).

Tudo indica que pY inibe a tradução principalmente na fase de alongação. Ela aparece quando o crescimento é impedido, ou por estresse ou pela fase estacionária, uma estratégia para sobrevivência em condições de estresse (Agafonov et al., 2001). Intrigante que em *C. pseudotuberculosis* 1002, a situação foi contrária ao observado em *E.coli*, onde a proteína representa uma ligação direta entre ribossomo e adaptação ao ambiente. Porque a bactéria estaria impedindo a síntese de proteínas em condição fisiológica normal de crescimento e no início da fase exponencial, tornando o resultado bastante interessante, mas difícil de esclarecer sem prévio conhecimento da sua regulação.

Em nível de comparação às outras condições, no meio submetido ao estresse ácido este gene foi induzido com *fold-change* de 1,38 e com RPKM de 6331,66 e no térmico foi também

reprimido em relação ao controle, que apresentou *fold-change* de 1,5 vezes a mais que o choque térmico. Se observarmos a figura 5, notamos que na condição de acidez (pH), houve menor número de células viáveis, se compararmos com as outras condições, em concordância com o trabalho de *E.coli*, mas deixando ainda uma dúvida em relação ao seu verdadeiro papel em *C. pseudotuberculosis* 1002.

O gene *mmdA*, codificante da subunidade alfa de metilmalonil-CoA descarboxilase, apresentou na condição controle um RPKM elevado em relação ao estresse osmótico, o qual foi reprimido quase 3 vezes em relação ao controle. No térmico foi reprimido, mas abaixo de duas vezes e na acidez foi induzido, mas com valor de RPKM próximo ao controle (anexo 8). A função desta enzima como bomba de sódio foi primeiramente relatado em *Veillonella parvula* (Hilpert e Dimroth, 1983). A enzima cataliza a única etapa de conservação de energia na rota de fermentação do succinato. A descarboxilação do (S)-metilmalonil-CoA é acoplada ao transporte de Na⁺ através da membrana, para fora da célula, criando assim uma força motriz de íons de sódio que é usada para síntese de ATP.

A metilmalonil-CoA descarboxilase é um complexo que consiste de cinco subunidades diferentes, em *Veillonella parvula*. A subunidade alfa foi caracterizada como carboxiltransferase (Huder e Dimroth, 1995). Em *C.pseudotuberculosis* 1002, tudo indica que o operon consiste de quatro subunidades e a subunidade alfa está abaixo do gene *pycB* (subunidade β do piruvato carboxilase). Este resultado, em *C.pseudotuberculosis* 1002, nos leva a supor que em condições fisiológicas normais, a expulsão do íon Na⁺ da célula parece ocorrer de forma mais acelerada, pois este gene é altamente expresso no controle, o que é indício de uma maior produção de ATP e assim, maior replicação.

Curioso, pois a célula no estresse osmótico, como estratégia de sobrevivência, acumula K⁺ e para que ocorra esta situação, a célula retira o Na⁺ do seu interior. Porém, justamente na condição de estresse, apesar do número de transcritos elevado, este gene se mostrou reprimido. Seria interessante estudar os reguladores deste gene para entender todo o processo de síntese de energia a partir da expulsão do Na⁺. É bem verdade que no controle, a célula apresenta um crescimento mais acelerado e possivelmente, graças ao aumento de síntese de energia. Logo, no estresse onde a síntese é menor em relação ao controle, apesar do ajuste para suportar a condição alterada, a bactéria cresce mais lentamente, confirmando a estratégia de reduzir o crescimento para garantir a permanência no meio alterado.

O gene *uspA*, codificante da suposta proteína de estresse universal, surpreendeu com os valores apresentados. A condição de estresse osmótico apresentou um valor de RPKM

considerável, mas em relação ao controle foi considerado reprimido. Analisando as outras condições, no choque térmico não foi significativo e na acidez também foi reprimido. Em *E.coli*, o gene demonstrou aumento na expressão em diversas condições de estresse, além do aumento na fase de crescimento estacionária. O gene demonstrou ainda ser dependente da ativação do promotor pelo $\sigma 70$ (Nyström e Neidhardt, 1992). Em *C.pseudotuberculosis* 1002, σA no estresse osmótico foi induzido em relação ao controle, o que surpreende ainda mais este resultado. Pelo valor de RPKM apresentado no estresse, não implica que a célula seja prejudicada pela condição que se encontra. Apesar de ter sido considerado reprimido em relação ao controle, apresentou número de transcritos elevado.

Tudo indica que este gene tenha uma importância em fase exponencial, mas em fase estacionária a contribuição deve ser ainda maior, tendo em vista o trabalho com *Salmonella typhimurium*, onde observaram uma resistência do microrganismo ao estresse oxidativo e para determinar se esta resistência era dependente da fase de crescimento, deixaram a linhagem selvagem e mutante para *uspA* exposta à fase exponencial e estacionária com H_2O_2 . Os resultados mostraram que na fase exponencial as células são mais sensíveis e reduzem a sobrevivência, mas a linhagem mutante não apresentou nenhuma resistência, levando a 100% de morte, o que seria indício do papel de UspA na resistência durante a fase exponencial. Na fase estacionária as células mutantes mostraram maior resistência, o que poderia ser também devido a atuação de outros genes auxiliando na sobrevivência, como catalase (Liu et al., 2007).

Este gene demonstrou contribuir para a patogenicidade de *Salmonella typhimurium* e apresentou importância na virulência pela resistência ao estresse universal. É fundamental ampliar o conhecimento dos mecanismos de sobrevivência ao estresse em bactéria, pois somente assim, pode-se desvendar os caminhos para reduzir a persistência do patógeno e as doenças por ele provocada (Liu et al., 2007).

6.3.3 Processos Biológicos do Estimulo Osmótico

6.3.3.1. Processo Induzido

A maior parte dos genes induzidos no estresse osmótico fazem parte do processo de metabolismo e biossíntese, indicando que a bactéria ainda no ambiente precário investe energia para sua replicação e sua sobrevivência. Observa-se tanto entre os induzidos (figura 17) quanto entre os reprimidos (figura 18), os processos metabólicos se destacando, como se a célula buscasse um equilíbrio para não ser prejudicada totalmente. Genes diferentes compõem os

processos entre induzidos e reprimidos (anexos 8 e 9, respectivamente), mas está certo que a célula apenas reduz o seu crescimento, não interrompe por completo e por isso consegue se manter no ambiente. Neste início da fase exponencial, a bactéria direciona também a resposta para síntese de macromoléculas, estocando energia e mantendo seu crescimento, como observamos nos genes envolvidos no processo biossintético.

O processo de transporte transmembrana está também em destaque. Os genes relacionados a este processo estão envolvidos na aquisição de ions (anexo 9) possivelmente buscando equilíbrio entre as moléculas para tentar minimizar os efeitos causados pelo ambiente hiperosmótico.

Entre os processos induzidos o que nos chama a atenção são aqueles menos representados, mas com grande significado, como o processo envolvido na adesão celular e sistema de dois componentes que são respostas específicas que irão contribuir ativamente na sobrevivência da bactéria no meio. O gene Cp1002_0988 o qual codifica uma proteína secretada com motivo I e o gene Cp1002_0989 que codifica proteína de membrana ancorada a superfície compõem o processo de adesão celular. Os dois genes estão presentes na ilha de patogenicidade com *fold-change* de 7 vezes a mais que o controle (Cp1002_0988) e quase 3 vezes a mais que o controle (Cp1002_0989) (tabela 17). Estas proteínas detectadas como secretadas demonstram ser importantes candidatos vacinais e outros estudos devem ser realizados para verificarmos esta possibilidade. No sistema de dois componentes os genes Cp1002_0431 (Two component LuxR family transcriptional) e Cp1002_0184 (em *C.pseudotuberculosis* 1002 anotado como *phoB*) aparecem contribuindo para estas funções e devem ser investigados como possíveis candidatos vacinais (anexo 9). Estes genes foram identificados como regulador de resposta do sistema de dois componentes ou seja, são osmosensores que regulam a expressão de genes que contribuem para o crescimento e sobrevivência em ambiente hostil.

6.3.3.2. Processo Reprimido

A resposta ao estresse é normalmente representada pela combinação da resposta específica, por minimizar os efeitos deletérios, e geral por reprimir genes relacionados a tradução e biogênese ribossomal (Hengge-Aronis, 2002), estando relacionados a paralisação ou redução do crescimento, buscando adaptação e ajuste da fisiologia em um novo ambiente (Jozefczuk, et al., 2010). No caso deste trabalho, como já citado acima, a bactéria não paraliza

sua replicação, apenas desacelera, pois a energia deve também ser distribuída para síntese de proteínas específicas à sua permanência no meio.

O processo de oxido-redução foi o mais representado, pois a respiração reduzindo, diminui também a probabilidade de produzir moléculas espécie oxigênio reativas. Outros processos envolvidos na regulação da transcrição, metabólico e processo do ciclo ácido tricarboxílico também podem ser visualizados, confirmando a redução no crescimento (figura 18).

Um resultado contraditório ao da literatura (Han et al 2005) foi o processo de transporte, presente entre os reprimidos, pois quando a bactéria é exposta ao ambiente hiperosmótico, a célula deve aumentar sua osmolaridade interna ou pela absorção ou sintetizando pequenas moléculas como potássio, glicina betaina, prolina, trealose (Meyer et al., 1990). É bem verdade que por estar em meio rico, a bactéria opte por não sintetizar a molécula e sim absorver do meio, pois não desperdiça energia, mas o que observamos é repressão do processo de transporte, composto pelos genes Cp1002_0623 (proteína transportadora do tipo ABC ligante ao soluto); Cp1002_0383 (*malk*) (ATPase transportadora de glicerol-3-fosfato) e Cp1002_1988 (transportador de nucleosídeo) (anexo 10) e sendo assim, provavelmente, nesta fase de crescimento *C. pseudotuberculosis* busca a solução em outros processos, e esta resposta via síntese de moléculas ou transporte, deve ocorrer em fase mais tardia.

6.4. Estimulon da acidez

6.4.1 Genes induzidos

Alguns microrganismos dependem de fatores que são encontrados em fagolisossomo, como sinais para acionar uma multiplicação intracelular, como *S. typhimurium* que requer um pH ácido para iniciar replicação intracelular e sintetizar fatores que permitem a persistência neste ambiente (Rathman et al., 1996). Outros, como micobactéria, parece modificar o fagossomo primário para escapar da acificação ou explorar a rota celular normal do hospedeiro para se proteger em vesícula celular com pH neutro (Sturgill-Koszycki et al., 1994). Pressupõem-se que outros patógenos usem sinais como presença de glicoproteínas lisossomais ou enzimas lisossomais para seletivamente expressar genes codificantes de proteínas que potencializam a sua sobrevivência e replicação diante de um ambiente fagolisossomal (Finlay e Falkow, 1997).

Corynebacterium pseudotuberculosis 1002, parece suportar o ambiente fagolisossomal, expressando genes que codificam proteínas que favoreçam sua permanência neste ambiente, possivelmente alterando o pH do meio, retirando/utilizando os íons que acidificam o meio. Entre os genes induzidos, *msrB* (20x), *dps* (8x), *lysR* (7x), chamaram a atenção pelo valor de *fold-change* e pelo valor de RPKM em relação ao controle (tabela 12).

O gene *msrB* supostamente codifica metionina sulfóxido redutase. A metionina (Met) é o aminoácido mais sensível à espécie oxigênio reativa e a oxidação do resíduo de Met na proteína (ou livre) resulta em uma estrutura de proteína alterada (ou impede a tradução) afetando drasticamente sua função. O resíduo de Met oxidado são reduzidos à metionina pelo sistema metionina sulfóxido redutase (Msr), que incluem principalmente MsrA e MsrB (Dhandayuthapani et al., 2009). A anotação em *C.pseudotuberculosis* 1002 demonstrou que MsrB é específica para redução de metionina-R-sulfóxido. A que tudo indica, em *C.pseudotuberculosis* 1002 apenas o domínio MsrB está presente. Patógenos deficientes em Msr demonstrou habilidade reduzida em aderir em células eucarióticas, sobreviver no hospedeiro e resistir ao estresse oxidativo, *in vitro* (Sasindran et al., 2007). Uma hipótese emergente é que os resíduos de Met em proteínas expostas na superfície atuam como “catadoras” para espécie oxigênio reativa em células e o sistema Msr reduz as metioninas oxidadas fornecendo um ciclo redox para remover a espécie oxigênio reativa do ambiente (Delaye et al., 2007).

Em *C.pseudotuberculosis* 1002, na condição onde simulamos a acidez dos macrófagos, apresentou alta expressão de gene *msrB*, sendo indício que o gene se torna ativo desde o início da fase exponencial e provavelmente exerça importante função nesta condição que intensifica o estresse oxidativo. Na realidade, os estudos revelam maior participação de MsrA na resistência ao estresse oxidativo e sobrevivência em macrófagos que MsrB, mas em *H.pylori* a contribuição foi confirmada, ainda que em nível menor que MsrA (Alamuri e Maier, 2004). Um indício da maior participação de MsrA poderia ser que a oxidação de metionina-R-sulfóxido seja menos frequente que a oxidação de metionina-S-sulfóxido, alvo de MsrA.

Contudo, a perda do domínio *msrA-msrB* ou somente de *msrB* em *H.pylori*, demonstrou redução na virulência em modelos murinos, provavelmente pela oxidação de proteínas importantes (Alamuri e Maier, 2004). O gene parece ser regulado por fator sigma de função extracitoplasmática em bactéria não patogênica, e apresenta forte indício que seja regulado também por fator sigma alternativo em bactérias patogênicas (Sasindran et al., 2007).

A indução do gene *msr* já foi relatada em outros ambientes, como a mudança do ambiente ácido para neutro em *Streptococcus gordonii* (Vriesema et al., 2000) na fase estacionária de *E.coli* (Moskovitz et al., 1995) e limitação de ferro em *H.pilory* (Alamuri e Maier, 2006). Em *C.pseudotuberculosis* 1002, a bactéria mostra-se “resistente” à condição de acidez, apresentando uma redução de cerca de 30% no crescimento e possivelmente seja pela contribuição deste gene.

O gene *dps* codifica a suposta proteína de proteção ao DNA durante inanição. Estas proteínas são capazes de estocar ferro na forma biodisponível e proteger as células contra estresse oxidativo. O grande problema enfrentado pelo microrganismo no hospedeiro é a baixa concentração de ferro livre na célula (concentração típica de 10^{-18} e 10^{-8} para Fe^{3+} e Fe^{2+} , respectivamente). Embora amplamente distribuído, sua tendência a oxidar formando o composto Fe^{3+} insolúvel pode romper a função celular normal. Conseqüentemente, ele se encontra complexado com diferentes proteínas. A presença de Fe^{2+} pode levar a geração de radicais hidroxila tóxicos de acordo com a reação de Fenton: $H_2O_2 + Fe^{2+} \rightarrow Fe^{3+} + OH^- + OH$ (Storz e Imlay, 1999).

O radical hidroxila é capaz de induzir quebra no DNA, peroxidação lipídica e degradação de biomoléculas. A proteção fornecida pela proteína Dps ocorre por duas formas: ou pela ligação ao íon Fe^{2+} evitando a formação de radicais hidroxila tóxicos catalizados pela reação de Fenton ou pela ligação ao DNA protegendo dos radicais oxidativos. A estrutura cristal de Dps revelou homologia estrutural com a ferritina (Haikarainen e Papageorgiou, 2010).

Em *C.pseudotuberculosis* 1002, o gene possivelmente apresentou maior expressão na condição de acidez, pois neste ambiente a produção de moléculas tendenciosas a produzir espécie oxigênio reativa é maior. Existe relatos demonstrando que em célula onde o pH diminui, a razão de HOO^- para O_2^- aumenta (Farr e Kogoma, 1991), e conseqüentemente a chance de produzir H_2O_2 também. Assim, a maior expressão do gene *dps* na acidez, em *C.pseudotuberculosis* 1002, sugere proteção desde o início da fase exponencial juntamente com a catalase que também se apresentou altamente expressa nesta condição. Este gene está relatado no item das ilhas de patogenicidade.

O gene *lysR*, codificante da possível proteína ativadora transcricional, se destacou pelo valor de RPKM apresentado na condição de estresse em relação ao controle (tabela 12). A família LysR representa o mais abundante e melhor caracterizado tipo de regulador nos procariotos. As pesquisas revelam que este regulador transcricional global atua ou como

ativador ou como repressor de diversos genes. Apresenta estrutura conservada com o motivo hélice-volta-hélice na extremidade N-terminal ligante ao DNA e a extremidade C-terminal domínio ligante ao co-indutor (Schell, 1993).

A conservação no genoma de bactérias extremamente diversificadas significa que estão envolvidos na regulação de genes com mesma forma em diversas funções, cujos produtos podem estar envolvidos no metabolismo, divisão celular, quorum sense, virulência, motilidade, fixação do nitrogênio, resposta ao estresse oxidativo, produção de toxina, ligação e secreção, entre outras. O gene *lysR* é transcrito quando a proteína LysR está dissociada do promotor. A proteína então, traduzida, liga acima do promotor do gene alvo, transcrito divergentemente. Quando o co-indutor interage com o regulador, a transcrição deste gene alvo é ativada (Maddocks e Oyston, 2008).

Os reguladores transcricionais globais são importantes no desenvolvimento de linhagens vacinais atenuadas e protetoras (Keller et al., 2008). A aplicação da pesquisa transcriptômica voltada aos estudos dos reguladores globais tem sido amplamente utilizada para desenvolvimento de vacinas, diagnóstico e desenvolvimento terapêutico (Maddocks e Oyston, 2008). Em *C.pseudotuberculosis* 1002, o domínio proteico apresentou homologia com o domínio OxyR, um regulador de resposta oxidativa em *E.coli* e *Salmonella typhimuricum* (Farr e Kogoma, 1991).

Todos os organismos aeróbicos devem se defender contra oxidantes tóxicos resultantes da redução incompleta do oxigênio (Schell, 1993). O OxyR não responde ao clássico co-indutor, mas sob mudança no redox pode alterar sua conformação e afinidade ao DNA. A oxidação da OxyR também influencia a interação com a RNA polimerase na região promotora, iniciando a transcrição (Zaim e Kierzek, 2003). Em *C.pseudotuberculosis* 1002 é provável que este gene contribua para a sobrevivência ao ambiente ácido que pode levar a formação de espécie oxigênio reativa. Pode estar envolvido com outras funções, mas outros estudos devem ser realizados, como desvendar o seu regulon, pois também pode contribuir na proteção contra LC.

6.4.2. Genes reprimidos

Como nas outras condições, várias proteínas hipotéticas foram identificadas, inclusive o gene mais reprimido, Cp1002_1888, codifica uma suposta proteína hipotética. Destacamos três genes que demonstraram um alto valor de RPKM no controle em relação ao estresse e o

fold-change foi acima de 3 vezes. O gene *odhI* apresentou-se 4x mais induzido no controle, *ppiA* apresentou 3,44 vezes e *pyrH* apresentou-se 3 vezes mais induzido no controle em relação ao estresse ácido (tabela 13).

A suposta proteína inibitória da atividade ODHc (complexo 2-oxoglutarate dehydrogenase, também conhecido como alfa cetoglutarase desidrogenase) é codificada pelo gene *odhI*. A proteína apresenta o domínio FHA (Forkhead associated domain), suposto domínio de sinalização nuclear, que pode ligar fosfotreonina, fosfoserina e algumas vezes fosfotirosina. Em eucariotos, proteínas contendo domínio FHA estão ligadas a regulação do ciclo celular, do reparo do DNA, ou da transcrição (Durocher et al., 2000). Em humanos, a enzima 2-oxoglutarate dehydrogenase é inativada por espécies oxidativas e toxinas (Bunik et al., 2007).

Esta enzima chave do ciclo ácido tricarboxílico mostrou ser regulada por uma enzima serina/protease kinase, a PknG. A inibição de ODH pode ocasionar distúrbio no metabolismo de energia. Esta conclusão foi obtida após mutarem o gene para PknG e observarem queda no metabolismo pela inibição de ODH, pois a PknG mostrou fosforilar OdhI na treonina 14, e por ser a ODH essencial para utilização de glutamina em *C.glutamicum*, assumiu-se que a falta de fosforilação, inibe sua atividade (Niebisch et al., 2006). Na acidez, em *C.pseudotuberculosis* 1002, a ODH então não foi tão prejudicada pelo seu inibidor, que foi reprimido.

Em *C.pseudotuberculosis* 1002, o gene *pknG* foi reprimido em todas as condições, mas não foi considerado significativo em nenhum estresse, e na acidez foi ainda mais reprimido com valor de RPKM menor em relação ao estresse osmótico e térmico. Como *pknG* foi reprimido, a probabilidade de fosforilar OdhI reduz, e, conseqüentemente o metabolismo seria prejudicado, pois a ODH neste caso seria inibida. Porém, o gene *odhI* foi reprimido, então menos genes estão disponíveis para codificar a proteína inibitória de ODH, deixando-a mais ativa e assim, o que possivelmente ocorreu foi uma inibição mais branda, visto que houve expressão do gene, e neste caso, como já descrito anteriormente, faz parte de uma estratégia de sobrevivência da bactéria, desacelerar o crescimento para adaptação em um novo meio (Jozefczuk et al., 2010), nos dando suporte para explicar a redução no número de células no meio com estresse.

Na forma não fosforilada, OdhI liga com alta afinidade na OdhA, uma suposta subunidade E1 do complexo alfa- cetoglutarase desidrogenase em *C.glutamicum* e assim inibe a atividade do complexo. Então, quando PknG fosforila OdhI, a OdhA é liberada (Niebisch et al., 2006). Em *C.pseudotuberculosis* 1002, o gene codificante da suposta proteína alfa-

ketoglutarase desidrogenase (OdhA), está anotado como *kgd*. Na condição de acidez este gene foi considerado induzido em relação ao controle, mas com *fold-change* de 1, 16 vezes, por isso não entrou nas nossas análises, cujo corte foi de 2x, mas confirma que a inibição da atividade não ocorreu por completo.

O gene *ppiA* codifica a suposta peptidyl-prolyl *cis-trans* isomerase A, também conhecida como ciclofilina, que é encontrada em todos os organismos estudados. A formação correta da ponte dissulfeto e a isomerização *cis-trans* de ligações peptídicas anterior a prolina são processos que determinam a taxa de dobramento de muitas proteínas. A etapa de conformação é lenta e pode ser acelerado por duas enzimas diferentes: proteína dissulfeto isomerase (PDI) e peptidyl-prolyl *cis-trans* isomerase (PPIases), respectivamente para formação correta da ponte e dobramento (Schönbrunner e Schmid, 1992).

Em *C.pseudotuberculosis* 1002, uma vez que o catalizador foi considerado reprimido em relação ao controle, nos infere que o processo de dobramento é mais lento no início da fase exponencial na condição de estresse, nas proteínas alvos desta enzima. É possível que seja uma enzima que apresenta atividade ótima em condições fisiológicas. Como em condição de estresse o crescimento é mais lento, é possível que as proteínas alvos desta enzima sejam aquelas envolvidas com crescimento ou sejam essenciais apenas em condições fisiológicas, onde requer produção mais acelerada. Contudo, são hipóteses que devem ser confirmadas experimentalmente.

Em *E.coli* a enzima está localizada no espaço extracitoplasmático e a inativação de enzimas PPIases não resultou em defeitos observáveis, mas alterou a taxa de enovelamento da proteína, mas qual papel estas enzimas exercem na maturação de proteínas no espaço extracitoplasmático não está claro. Outros papéis para cicloporina incluem função chaperona e sinalização celular, contudo não tem relato destas funções para PPIA (Justice et al., 2005).

Seria interessante revelar os substratos desta enzima, conhecer seus alvos para descobrir qual a influência desta isomerase no dobramento e maturação de proteínas e onde estão localizadas estas proteínas em *C.pseudotuberculosis* 1002, além de revelar o (s) gene (s) que a regula (m).

Outro gene que apresentou uma diferença interessante entre os RPKMs do estresse e controle foi o *pyrH*. Este gene codifica a uridylate kinase, uma enzima que representa uma classe particular de proteínas nucleosídeo monofosfato quinase (NMP), que são enzimas essenciais no metabolismo de nucleotídeos, mas a uridylate kinase que pode ainda estar envolvida com a divisão celular, não mostra similaridade com nenhuma NMP quinase

conhecida e está sujeita a mecanismos regulatórios complexos. O gene *pyrH* é considerado essencial e por isso o produto deste gene pode ser um interessante alvo para drogas antibacterianas e pode exercer outras funções além de catálise. Em *E.coli*, esta enzima se mostrou localizada próximo das membranas, mas outros experimentos precisam ser realizados para certificar a descoberta (Landais et al., 1999)

Mutantes dos genes *luxS* e *smcR* de *Vibrio vulnificus*, bactéria Gram-negativa, demonstraram reprimir 8 genes codificantes de proteínas, incluindo uridylate kinase, envolvidas na virulência, adaptação ao ambiente com estresse, biossíntese de LPS e multiplicação celular (Shin et al., 2007). A ausência da função normal de UMP kinase causa defeitos no sítio de septação, comprometendo a divisão celular. Como estão envolvidas na síntese de trifosfato nucleosídeo, precursor das principais moléculas de RNA, DNA e fosfolípidos, em *C.pseudotuberculosis* 1002 a produção de novas moléculas foi reduzida, o crescimento foi mais lento, confirmando novamente a estratégia de reduzir o crescimento para sobreviver ao meio hostil.

6.4.3. Processo Biológico do Estimulo no meio ácido

6.4.3.1. Processo Induzido

Na acidez, genes regulatórios e da família sigma representam o processo de regulação da transcrição (anexo 11), o mais destacado entre os vinte processos mais induzidos, representado por 10 genes (figura 19). Isto implica que na acidez, um maior número de genes em resposta ao estresse está presente, uma vez que a maioria destes reguladores são ativos por estímulo ambiental (Perera e Grove, 2011) (Schell, 1993) (Gao et al., 2011). Estes reguladores são alvos importantes como candidatos vacinais e merecem atenção em estudos futuros.

Um dos genes, Cp1002_0718, codifica um suposto regulador transcricional da família AraC, que regula diversas funções que partem desde metabolismo do carbono até resposta ao estresse associada a virulência. O mutante para este gene em *M. tuberculosis* reduziu sua sobrevivência em macrófagos, sugerindo que a proteína regula genes importantes para a virulência (Frota et al., 2004).

O gene Cp1002_1682, codifica o regulador transcricional da família ArsR, o qual, em *M.tuberculosis*, percebe e responde ao estresse originado por metal e modula a transcrição de genes envolvidos na absorção de metal, efluxo, sequestro ou detoxificação. Normalmente, a maioria dos reguladores da família ArsR atuam como repressores da expressão dos genes e liberam-se do promotor quando há abundância de metal na célula, permitindo então a

expressão dos genes. No entanto, ArsR demonstrou regular positivamente o gene *phoP*, sugerindo influenciar a patogenicidade pela regulação do operon *phoPR*, uma vez que *phoPR* é essencial para a virulência e exerce importante papel na patogênese e em uma linhagem avirulenta foi observado um polimorfismo no sítio ligante do operon, onde prejudicou a ligação do regulador (Gao et al., 2011).

Diferente do estresse osmótico, 4 genes compõem o processo de adesão celular, representados pelos genes Cp1002_0023 – proteína secretada com motivo I; Cp1002_1083- proteína secretada com motivo I; Cp1002_1964- proteína secretada com motivo I (em *C.pseudotuberculosis* 1002 todas anotadas como proteína hipotética) e Cp1002_0439- superfamília do tipo ABC- proteína ligante ao substrato), o que sugere mais uma vez, que a condição de acidez que mais contribuiu para o aparecimento de genes relacionados a virulência, como também o processo de transporte de ferro com alta afinidade composto pelos genes Cp1002_0289 –proteína ligante periplasmática (em *C.pseudotuberculosis* 1002 anotado como suposto transportador do tipo ABC) e Cp1002_0031-proteína ligante ao ferro-sideróforo (*fagD*), além do processo envolvido na patogênese, representado pelo gene Cp1002_0027- fosfolipase D (anexo 11), que são processos de extrema importância para o início de um processo de infecção, que merece atenção para estudos futuros quanto ao envolvimento no desencadeamento da doença e por isso devem ser avaliados quanto a possibilidade de se tornarem alvos vacinais.

6.4.3.2. Processo Reprimido

Novamente, entre os genes reprimidos (figura 20) estão aqueles envolvidos com o crescimento bacteriano ou biossíntese. Interessante que o processo de regulação da transcrição está entre os mais induzidos e entre os mais reprimidos na célula. Para persistir no ambiente do hospedeiro, *M.tuberculosis* entra em estado de dormência e rapidamente reprime a biogênese do ribossomo para combinar com a necessidade reduzida de tradução, uma resposta que requer coordenada regulação da transcrição de todos os componentes do ribossomo (Betts et al., 2002).

O gene Cp1002_1761 (*carD*) (anexo 12) foi identificado como uma nova proteína de interação à RNA polimerase que controla a transcrição do rRNA e genes de proteínas ribossomais em micobactérias tanto no estado estacionário quanto durante diversos estresses celulares, incluindo inanição, danos ao DNA, estresse oxidativo. Esta proteína CarD é membro de uma família que modula a transcrição sem ligação ao DNA, como parte da

resposta global ao estresse em procarioto. Demonstrou ser essencial para a viabilidade de micobacterias sob todas as condições (Stallings et al., 2009). Em *C.pseudotuberculosis* 1002, o gene foi reprimido na acidez, indicando menor importância nesta fase de crescimento, visto que as proteínas ribossomais foram induzidas na condição de estresse ácido (tabela 8) e como em *M.tuberculosis*, pode ser que atue em fase mais tardia de crescimento. Como dito anteriormente, a bactéria desacelera seu crescimento, mas não paraliza por completo e possivelmente por isso, o gene não foi induzido, pois do contrário poderia levar a paralização completa do crescimento neste início de fase exponencial.

No processo de transporte o gene Cp1002_1670 (*ctaD*) (anexo 12) que codifica citocromo c oxidase subunidade I, está entre os representantes e esta enzima está envolvida no transporte de moléculas, íon, ou elétron através da membrana. A citocromo c oxidase é uma enzima chave na respiração celular, membro terminal da cadeia de transporte de elétron onde uma série de reação de redox mediada pela transferência de elétron resulta em energia livre usada para formação de ATP e NAD (Hannappel et al., 2011). Como o processo foi reprimido, é indício que a respiração foi reduzida e conseqüentemente a produção de ATP e NAD também, confirmando retardo no crescimento.

6.5. Estimulon do choque térmico

6.5.1. Genes induzidos

No estresse térmico o gene que apresentou um valor de RPKM bem acima do controle foi o *hspR*, cujo *fold-chance* foi de 4,49 vezes (tabela 14). Os valores de RPKM dos genes em geral, não estão altos nesta condição como nos outros estresses relatados e a quantidade de genes obtidos foi menor, depois de considerarmos o valor de corte selecionado. Podemos supor que o número menor de genes representados nesta condição, pode ser devido a uma dependência maior pela densidade celular, pois quando fez o corte, o número de genes diferencialmente expresso reduziu bastante. Dos 26 genes com *fold-change* ≥ 2 , 30% se referem às proteínas hipotéticas (tabela 14).

O gene *hspR* codifica um suposto regulador transcricional de chaperona. As células bacterianas sob elevada temperatura, acumulam proteínas agregadas, parcialmente degradadas e proteínas com falha no enovelamento, assim, como resposta, aumenta a produção de fatores protetores nomeados como proteínas do choque térmico (HSPs), que funcionam como proteases e chaperonas aliviando os efeitos das proteínas aberrantes (Gross, 1996). Porém,

muitas HSPs exercem papel chave no metabolismo celular sob todas as condições de crescimento, monitorando o enovelamento, montagem e translocação de proteínas celulares (Craig et al., 1993).

Em *H.pylori*, existem operons regulados pela HspR na condição de estresse osmótico (Spohn e Scarlato, 1999), e em *C.pseudotuberculosis* 1002, o gene na verdade foi induzido nas três condições, porém, com o *fold change* acima de 2, foi somente no estresse térmico. As respostas por choque térmico em bactérias não são limitadas a mudanças de temperatura e normalmente é uma resposta geral aos estresses. Por isso, muitas HSPs são induzidas por outras alterações ambientais, como adição de etanol, metais pesados, alta osmolaridade, alteração no pH, inanição, estresse oxidativo ou interação com hospedeiro (Gophna e Ron, 2003), pois qualquer ambiente com estresse pode levar à ocorrência de proteínas aberrantes.

A indução de genes envolvidos na resposta ao estresse térmico, provavelmente afeta a patogenicidade indiretamente pelo aumento da resistência bacteriana às defesas do hospedeiro ou regulando genes de virulência (Gophna e Ron, 2003). A regulação para indução dos genes é complexa e diversa entre os procariotos, envolvendo mecanismos regulatórios positivos e negativos. A HspR liga na sequência HAIR (repetição invertida associada a HspR), conferindo a repressão dos promotores responsáveis pelo controle de HSPs na ausência do estresse (Servant e Mazodier, 2001).

Em *Mycobacterium*, a HspR reprime o operon *dnaK-grpE-dnaJ-hspR*, seu próprio gene, ligando na região HAIR localizada na região 5'UTR destes genes (Gupta, Bandyopadhyay, Gupta, 2008). Nesta mesma bactéria o gene parece regular outros genes, além do operon, ligados a virulência/patogenicidade. A própria HspR pode ser um fator de virulência, uma vez que o mutante resultou em perda da sua persistência (Stewart et al., 2002) e a chaperona DnaK, que é altamente antigênica, torna-se super expressa em mutantes para *hspR*, resultando em melhor resposta imune no hospedeiro (Gupta, Bandyopadhyay, Gupta, 2008).

A ligação de HspR na região HAIR não é sensível ao calor e parece ser dependente de proteína renaturada *in vitro* (Stewart et al., 2001) e tem sido sugerido que a presença de outros fatores podem exercer o papel de conferir a sensibilidade a este ambiente. Em *Streptomyces coelicolor* (Bucca et al., 2003) e em *Mycobacterium* (Gupta, Bandyopadhyay, Gupta, 2008) demonstraram que a DnaK atua também como co-repressor para HspR. Logo, é possível que reguladores como HspR sejam controlados por múltiplos fatores, um dos quais certamente é a DnaK. Em resposta ao estresse térmico, ou outro estresse, espera-se que HspR seja desnaturada e liberada do operador HAIR. No entanto a célula deve se reestabelecer o mais

breve possível, e para tal, a HspR tem que voltar a sua forma nativa rapidamente, e neste contexto, a presença de GroEL junto com DnaK pode ser um fator importante (Das Gupta et al., 2008).

Em *H.pylori*, os operons *groESL*, *hrcA-grpE-dnaK* e *cbpA-hspR-orf*, codificantes das principais chaperonas, são transcritos pelo fator de resposta vegetativa, sigma 80 e são regulados negativamente pelo repressor HspR, uma vez que o nível de mRNA destes operons aumentou 8 vezes na linhagem mutante para *hspR*. Outros experimentos revelaram que HspR liga diretamente no promotor destes três operons (Spohn e Scarlato, 1999).

Em *Streptomyces*, a síntese de HSPs, como DnaK, ClpB, GroEL e HSP18 é controlada negativamente em nível transcricional por pelo menos três diferentes repressores, sendo que a HspR é responsável pelo controle de *dnaK* e *clpB* (Servant e Mazodier, 2001), as outras proteínas são reguladas por outros reguladores.

Em *C.pseudotuberculosis* 1002, o gene *hspR* está localizado na fita reversa, abaixo dos genes *dnaJ*, *grpE* e *dnaK*, o que seria indício de serem regulados pela HspR, porém estudos para confirmação da regulação destes genes pelo regulador HspR serão necessários. Todos os três genes foram induzidos no estresse térmico com *fold-change* acima de dois (anexo 8), porém não entraram nesta análise como únicos do estresse térmico porque estes genes apresentaram um *fold-change* ainda maior na condição de acidez. O gene *dnaK* não se encontra no anexo 8, pois não foi considerado significativo na condição de estresse osmótico e neste anexo visualiza-se genes que foram considerados como significativos nas três condições ao mesmo tempo, mas verificando seu RPKM, apresentou um valor de 1211,46 no controle e RPKM de 3173,01 no estresse térmico, um valor bem acima do RPKM encontrado para *hspR*.

O gene *dnaJ* é conhecido por facilitar a interação de DnaK com o substrato (Liberek et al., 1995) e o *grpE* é considerado um fator de troca de nucleotídeo que substitui o ADP ligado à DnaK por ATP. O ATP ligado à DnaK tem menor afinidade pelo substrato (Brehmer et al., 2004). É possível que em *C.pseudotuberculosis* 1002 o operon seja formado por *dnak-grpE-dnaJ-hspR*, pelas posições que ocupam no genoma. Assim, sob alta temperatura a proteína HspR se dissocia do operador ativando o operon. O transcrito *hspR* que parece fazer parte do operon (figura 21) é induzido obviamente. Se a célula voltasse para o ambiente fisiológico normal, a proteína HspR retornaria a sua forma nativa e a DnaK que atua também como co-repressora, se ligaria nesta proteína, promovendo a ligação novamente à sequência HAIR, reprimindo o operon. Uma hipótese para atuação do co-repressor DnaK seria o baixo nível de

DnaK, como sinal de que o ambiente já estaria normalizado, e assim, esta proteína deve-se então ligar à HspR, para então inibir a transcrição.

Contudo, como relatado anteriormente, o mecanismo regulatório da síntese de HSPs é complexo e difere amplamente entre os organismos, por isso estudos específicos relacionados ao regulador *hspR*, nesta bactéria, seriam necessários uma vez que, genes de resposta ao estresse térmico parecem estar envolvidos na virulência e contribuir para a sobrevivência da bactéria em ambiente hostil.

Entre os induzidos deste estresse, outro gene que apresentou um RPKM com um valor interessante em relação ao controle foi o *ahpD*, o qual codifica *Alkyl hydroperoxide reductase*. Esta enzima juntamente com a AhpC constituem enzimas detoxificante, importantes para proteção contra metabólitos oxidativos e nítrico reativos e são controladas pelo regulador OxyR (Olsen et al., 2000). As proteínas foram codificadas em grande quantidade em *Mycobacterium avium* subsp. *Paratuberculosis* e não foi detectada por Western Blot em outras micobactérias e em outras espécies, inclusive *C.pseudotuberculosis*, exceto em menor quantidade em *M. gordonae* quando cresceu sem exposição ao estresse oxidativo. A alta quantidade detectada das proteínas AhpC e AhpD sem necessidade de indução por peróxido é uma característica única desta bactéria (Olsen et al., 2000).

No trabalho acima, a resposta foi produzida pelos dois antígenos, principalmente por AhpC, indicando que estas proteínas podem exercer papel de proteção contra paratuberculosis. A resposta gerada contra AhpC e AhpD mostrou que estas proteínas podem potencialmente ser usadas no diagnóstico de paratuberculosis com um teste baseado em IFN- γ , e podem se tornar bons candidatos para inclusão em novas vacinas contra esta doença (Olsen et al., 2000).

Em *C.pseudotuberculosis* 1002, o gene *ahpD*, dependente de estresse para indução, pode ser tornar bom candidato, uma vez que foi induzido em todas as condições, mas somente a 50°C mostrou um *fold-change* acima de dois. Não tem relatos da influência de choque térmico na expressão deste gene, possivelmente exerce outras funções de proteção além das que são apresentadas. O gene *lysR* (homólogo ao *oxyR*) está localizado próximo aos dois gene, porém em fitas opostas, indício de uma possível regulação por este regulador, em concordância com os dados de Olsen et al (2000), contudo apesar de ter sido considerado significativo e induzido, a expressão está muito similar ao controle com *fold_change* pouco acima de 1.

6.5.2. Genes reprimidos

Apenas seis genes foram considerados reprimidos com expressão pelo menos duas vezes menor em relação ao controle, sendo duas proteínas hipotéticas (tabela 15).

O gene *atpC* apresentou um RPKM muito alto no controle e codifica F₀F₁ ATP synthase subunidade épsilon, que é um tipo de sistema de transporte de moléculas e íons dependentes de ATP. Esta enzima cataliza a síntese de ATP utilizando potencial eletroquímico de prótons (ou íons de sódio em algumas bactérias) através da membrana. Esta enzima também funciona na direção reversa, quando o potencial se torna ineficiente: ela cataliza a bomba de próton para produzir um potencial eletroquímico para hidrolizar ATP e formar ADP e Pi (fosfato inorgânico).

Consiste de dois motores rotativos: O domínio F₀ (transportador de próton) que é integral da membrana e gera a força sobre a translocação de prótons que é impulsionado por potencial eletroquímico de prótons e composto por 3 subunidades (a, b e c) e domínio F₁ (contém o sitio catalítico) que é a parte solúvel em água, que gira sua subunidade interna para hidrolisar ATP e é composto por 5 subunidades: $\alpha_3\beta_3\gamma\delta\epsilon$ (Okuno et al., 2011). Em suma, as ATPases do tipo F utilizam um gradiente eletroquímico para sintetizar ATP ou hidrolizar ATP para reverter o gradiente (Kato-Yamada et al., 2000)

As ATPases do tipo F tem papel fundamental na manutenção do equilíbrio do pH e na resposta e adaptação a diferentes estresses e em *Streptococcus pneumonia* sua atividade foi aumentada em condições ácidas, assim como seu mRNA aumentou, ou o mRNA apresentou maior estabilidade neste ambiente (Martín-Galiano et al., 2001). Para comparação, em *C.pseudotuberculosis* 1002 na condição de acidez foi considerado induzido, porém com *fold-change* um pouco acima de 1. Como se acredita na influência do choque ácido na estabilidade do mRNA, no estresse térmico podemos supor que a alta temperatura tenha prejudicado a estabilidade do mRNA como uma das justificativas de sua repressão. Outros estudos indicaram que a expressão do cluster de *atp* é mais correlacionado com a taxa de crescimento do que com o valor de pH do meio (Barriuso-Iglesias et al., 2006) (Jakob et al., 2007)

Em bactérias com cadeia respiratória, o papel primário da enzima é para sintetizar o ATP a partir de um gradiente de próton proveniente desta cadeia. Ao contrário, a bactéria que não tem a cadeia respiratória, seu papel é criar um gradiente de próton com energia fornecida pela hidrólise de ATP e manter o pH intracelular via extrusão do próton (Futai, 1989)

Em eucariotos, existem vários sistemas regulatórios para o complexo F_0F_1 , como a F_0F_1 mitocondrial que tem uma proteína inibitória específica de ATPase e suas proteínas auxiliares, onde acredita-se manter o potencial eletroquímico de prótons ou o nível de ATP evitando a síntese de ATP/atividade de hidrólise de F_0F_1 ATPase, quando a atividade respiratória é reduzida. Em *E.coli*, a subunidade ϵ de F_0F_1 é o único inibidor endógeno da atividade ATPase (Kato et al., 1997).

Quando a concentração celular de ATP é reduzida e a célula sofre com inanição, o sistema regulatório está ativo, onde a subunidade ϵ inibe o sistema de hidrólise, mas não a síntese de ATP (Feniouk e Junge, 2005). Logo, quando a célula se depara com estas situações ela tem que evitar a ação de F_0F_1 ATPase e salvar ATP (Kato-Yamada et al., 1999). Esta função inibitória acredita-se ser fisiologicamente importante para evitar o consumo de ATP (Okuno, Lino, Noji, 2011). Em estudo com *Bacillus* termófilos, o efeito inibitório da subunidade ϵ é observado somente na fase inicial de hidrólise de ATP em baixa concentração de ATP como em período lag muito longo (Kato et al., 1997).

Em *C.pseudotuberculosis* 1002 o que aconteceu foi redução na transcrição do gene que possivelmente codifica a subunidade ϵ do complexo F_1 , o que seria indício que a atividade da F_0F_1 ATPase não foi interrompida, e consequentemente a hidrólise de ATP não foi interrompida. Curioso, pois a respiração estava reduzida, o crescimento estava mais lento, a síntese estaria reduzida. Para verificar este fato, observamos as subunidades responsáveis por esta ação. A subunidade β (*atpD*) que catalisa a reação de síntese do ATP (Dimroth et al., 2000) foi reprimida nesta condição, confirmando que a síntese está realmente menor, porém não entrou na análise porque não foi considerado significativo. É o movimento da subunidade γ (*atpG*) que providencia a energia necessária para os centros ativos das subunidades β sofrerem alterações que permitam a produção e libertação de ATP (Dimroth et al., 2000) cujo gene foi também reprimido, mas com *fold-change* fora do corte estabelecido por isso não entrou na análise.

O que parece é que mesmo estando a síntese reduzida, a hidrólise não foi interrompida. A medida que foi produzindo, foi consumindo. Não houve acúmulo de ATP, contradizendo outros autores citados acima (Kato-Yamada et al., 1999; Okuno, Lino, Noji, 2011). Porém, em *Bacillus* a inibição ocorreu somente na fase lag prolongada, assim, nossos resultados estariam de acordo com Kato et al (1997), pois a extração foi realizada no início da fase exponencial.

A necessidade por energia é verdadeira; a célula precisava de recursos para continuar seu desenvolvimento, assim, ao hidrolisar o ATP, produziria energia para síntese de outras

moléculas. Justamente pela queda na síntese, o desenvolvimento acaba sendo prejudicado. Como a bactéria se encontra em ambiente precário, precisa sintetizar proteínas que sejam necessárias para a sua sobrevivência (específicas) e para tal, o processo catabólico deveria estar ativo, pois assim, produziria energia para reações endoenergéticas.

6.5.3. Processo Biológico do Estimulo no meio térmico

6.5.3.1. Processo Induzido

Entre os induzidos no estresse térmico, o que chamou atenção foi o processo biossintético glicina betaina (figura 22) que era esperado encontrar no estresse osmótico. O gene que compõe este processo é o Cp1002_1785 (*betA*) o qual codifica colina desidrogenase que foi induzido com *fold-change* de 5 vezes a mais que o controle, como pode ser observado na tabela 14. Esta proteína pertence a família de oxidoreductase, que catalisa a oxidação de colina para glicina betaina via aldeído betaina como intermediário. Promove maior tolerância à hipersalinidade e congelamento, além de contribuir ao equilíbrio osmótico da célula sob estresse. Os osmoprotetores não desempenham papel somente de equilíbrio osmótico, mas atuam também como estabilizadores efetivos da função enzimática, fornecendo proteção contra salinidade, altas temperaturas, congelamento, descongelamento e até mesmo ao seco (Rajan et al., 2010).

O processo de transporte induzido neste ambiente está representado pelo gene Cp1002_1784 (*betT*), (anexo 13) que codifica proteína de transporte por alta afinidade de colina, acentuando ainda mais a participação desta molécula na resistência ao estresse. BetT é um transportador secundário dependente do sódio (Tondervik e Strom, 2007).

O sistema Bet (*betTIBA*) é responsável pela síntese de glicina betaina a partir da colina presente externamente. Em *E.coli*, o regulon Bet compreende os genes estruturais *betT*, codificante do transportador de colina, *betA* codificante da colina desidrogenase, *betB* codificante da betaina aldeído desidrogenase e o gene *betI* codifica um ligante ao DNA e regulador sensor de colina (Tondervik e Strom, 2007). Em *C. pseudotuberculosis* 1002 somente *betT* e *betA* estão presentes.

Processos relacionados à resposta ao estresse e regulação da transcrição foram representados pelo gene Cp1002_1895 codificantes da proteína regulatória de choque térmico (HspR), já relatada no tópico Estimulon do choque térmico induzido.

O processo de adesão também presente neste ambiente foi composto pelo gene Cp1002_0014 que codifica uma proteína hipotética e segundo Blast2GO é uma proteína secretada (anexo 13) e pode se tornar importante candidato vacinal.

6.5.3.2. Processo Reprimido

Poucos genes foram selecionados entre os reprimidos do estresse térmico (figura 23), mas o resultado já indica processos envolvidos com biossíntese e metabolismo e assim, como já descrito no item Estimulon do choque térmico reprimido, menor probabilidade de estocar energia levando ao retardo no crescimento (anexo 14).

6.6. Genes codificantes dos fatores sigmas

No seu ambiente natural, ou no hospedeiro durante o processo de infecção, a bactéria enfrenta perturbações que exige uma resposta rápida e adaptativa para garantir sua sobrevivência. A resposta é acompanhada por alterações no padrão de expressão de genes cujos produtos são requeridos para combater, diretamente, a natureza deletéria do estresse (Hecker e Volker 2001). A indução da transcrição de genes que respondem ao estresse é alcançada pela ativação de fatores de transcrição que interagem com RNA polimerase para expressão gênica coordenada (Marles-Wright e Lewis, 2007).

Uma família de fatores de transcrição com o papel na resistência ao estresse é a subunidade da RNA polimerase, o fator sigma, que é essencial para o início da transcrição por exercer papel principal no reconhecimento do promotor (Haldenwang, 1995). Cada um dos vários fatores sigmas da célula é requerido para a transcrição de subconjuntos específicos de genes/operon dentro do seu *regulon* (Mooney, Darst, Landick, 2005). Regulon inclui todos os alvos controlados diretamente ou indiretamente por um único fator de transcrição (Cases e De Lorenzo, 2005).

O genoma de *C.pseudotuberculosis* 1002 contém oito genes codificantes de fatores sigmas, diferente de *C.glutamicum*, não patogênica, que possui sete sigmas sugerindo que o *sigK* a mais em *C. pseudotuberculosis* possa estar envolvido em algum mecanismo de virulência nesta bactéria. Os sigmas incluem fator sigma essencial SigA, não essencial SigB e

seis fatores alternativos, SigC, SigD, SigE, SigH, SigK, SigM, que possivelmente pertencem ao grupo dos fatores extracitoplasmáticos (Kalinowski et al., 2003) que são responsáveis pela regulação de genes envolvidos no transporte, adaptação da parede celular, ou secreção dentro do periplasma (Gram-negativas) ou ambiente extracelular (Kazmierczak, Wiedmann, Boor, 2005).

Na figura 24 pode-se observar a diferença de expressão de cada um destes genes em cada condição testada. Os genes *sigA* e *sigH* codificantes do fator sigma RpoD e do SigH, respectivamente, foram significativos nas três condições. Os outros foram significativos em uma ou duas condições. Seis genes (*sigA*, *sigB*, *sigC*, *sigD*, *sigE*, *sigH*) apresentaram um valor alto referente ao número de transcritos sequenciados representando a região ativa no genoma. Somente os genes *sigK* e *sigM*, apresentaram um valor baixo, se levarmos em consideração os valores dos outros genes deste item (tabela 16).

O *sigA* foi considerado induzido no estresse osmótico e temperatura, e reprimido na condição de acidez em relação ao controle, mas apresentou um valor de RPKM bem próximo ao controle, nesta última condição (tabela 16). Interessante que em *C. glutamicum* Larisch et al (2007) observaram que na transição da fase de crescimento exponencial para a estacionária, a transcrição do *sigB* é aumentada, no mesmo instante que a expressão de *sigA* é reduzida. Em *C.pseudotuberculosis* 1002, pode-se observar que na condição onde a expressão de *sigB* foi aumentada, *sigA* foi reduzida, e vice-versa. Este resultado indica que esta situação observada pelos referidos autores ocorre já no início da fase exponencial em *C.pseudotuberculosis* 1002.

O *sigA* codifica possivelmente, o fator sigma RpoD (sigma A) ou como nome alternativo, sigma 70. Este fator promove a ligação da RNA polimerase a sítios específicos ativando a transcrição da maioria dos genes essenciais ao crescimento exponencial de *Escherichia coli* (Helmann e Chamberlin, 1988). Em *C.pseudotuberculosis* 1002, este gene apresentou os quatro domínios conservados presentes na família do sigma 70. Toda bactéria apresenta este fator sigma primário, que é essencial e requerido para a maioria dos genes. Além dos fatores sigmas, existem também os fatores sigmas alternativos, que são dispensáveis e frequentemente induzidos em resposta a condições específicas, em resposta ao estresse por exemplo. Alguns destes fatores estão envolvidos na regulação da expressão de genes de virulência (Gomez et al, 1998).

Os sigmas primários e a maioria dos alternativos pertencem a família definida por sigma70 em *E.coli*, onde quatro regiões conservadas têm sido identificadas em sua sequência

de aminoácido (Lonetto et al., 1992). Em *M. smegmatis*, o *sigA* demonstrou ser fator primário, essencial e transcrito na fase exponencial e estacionária, sendo maior na fase estacionária.

No trabalho de Castro (Dissertação, 2009), trabalhando com PCR em tempo real, foi aplicado estresse oxidativo proveniente de peróxido de hidrogênio em *C.pseudotuberculosis* 1002 no início da fase exponencial ($DO_{600nm}=0,2$) e após 15 min, observou que o *sigA* apresentou aumento na expressão, porém não foi considerada uma diferença significativa em relação ao controle.

Alternativamente, σA pode ser um fator sigma alternativo especificamente requerido para expressão de genes de virulência. (Gomez et al., 1998). Um mecanismo similar foi sugerido em *E.coli*, onde a indução do gene *rpoD*, codificante do $\sigma 70$, pela falta de aminoácido e choque térmico demonstrou ser mecanismo de recuperação a partir do estresse (Taylor et al., 1984). Em *C.pseudotuberculosis* 1002, como *sigA* mostrou alteração na expressão entre controle e estresses, pode ser que esteja exercendo também, sua função de fator sigma alternativo.

O *sigB* apresentou homologia significativa ao fator primário, mas não foi considerado essencial em *M. smegmatis* (Gomez et al., 1998). Trata-se de um fator sigma alternativo de resposta ao estresse e em *Listeria monocytogenes* o σB contribuiu para sobrevivência celular sob diferentes condições de estresses (Ferreira, O'Byrne, Boor, 2001; Ferreira et al., 2003). A ausência de σB reduziu a habilidade desta espécie em invadir células epiteliais (Garner et al., 2006), bem como sua virulência em murinos (Nadon et al., 2002). O fator σB contribui para virulência em diversos patógenos gram-positivos (Kazmierczak, Wiedmann, Boor, 2005). O regulon de σB contém pelo menos 127 genes, incluindo aqueles com função de resistência ao estresse, regulação transcricional e transporte de membrana (Price et al., 2001). Em *B. subtilis* e *L. monocytogenes*, o *sigB* faz parte de um operon contendo 8 genes. Apenas duas proteínas (RsbV e RsbW) deste operon apresentaram conservação entre diferentes espécies examinadas e assim, parece o mínimo essencial para a regulação da atividade de σB (Ferreira et al., 2004).

Em *C.pseudotuberculosis* 1002 este gene só foi considerado significativo na condição de acidez e foi altamente induzido, com uma razão de expressão três vezes maior que no controle (tabela 16 e figura 24). No genoma de *C.pseudotuberculosis* não tem anotado o operon do qual faria parte o σB , nem pelo menos as duas proteínas essenciais a regulação do σB . O gene *sigB* está localizado entre o gene *dtd* e *dtxR*. Nenhum anti-sigma próximo foi localizado, por estudo *in silico*. A organização no genoma de *C.pseudotuberculosis* 1002 está bem diferente do que se encontra na literatura a respeito do *sigB* em outras espécies como *B.*

subtilis e *L. monocytogenes*, porém similar a *M. tuberculosis*, onde *mysB*, codificante do fator sigma B está localizado a 3kb (em *C.pseudotuberculosis* 1002 5kb) abaixo do *sigA*, e não é flanqueado por genes codificantes de proteínas regulatórias do fator sigma (Doukhan, et al., 1995) e como em *C.pseudotuberculosis* 1002, o gene *dtxR* está abaixo do gene *mysB*.

Intrigante é que o gene *dtxR* foi induzido somente na condição de acidez, como o *sigB*. Em *Corynebacterium diphtheria* conseguiram identificar dois promotores que determinam a transcrição de *dtxR*. Um promotor localizado acima do *sigB*, que parece ser controlado por um fator sigma de função extracitoplasmática (ECF) e o outro promotor localizado na região intergênica entre *sigB* e *dtxR*, similar ao promotor utilizado pelo fator sigma primário em outras espécies de actinomicetos (Oram, Jacobson, Holme, 2006). Ainda neste trabalho, observaram que o aumento da transcrição de *dtxR* devido à exposição a acidez e frio, pode ser devido ao aumento de 5 ou 6 vezes o aumento da transcrição em *sigB* nestas condições. Em adição, o número de transcritos contendo *sigB* foi induzido em diferentes condições ambientais, incluindo exposição a acidez, frio, alta temperatura, etanol, ou duodecil sulfato de sódio e a exposição ao sal ou peróxido de hidrogênio não afetou a transcrição do *sigB*. O que sugere que o promotor acima do *sigB* possa ser reconhecido pelo fator sigma de função extracitoplasmática (Oram, Jacobson, Holme, 2006).

A expressão do *sigC* foi considerada significativa apenas na condição de estresse ácido, com um aumento de quase duas vezes em relação ao controle (tabela 16 e figura 24). Na condição de estresse osmótico e temperatura foi reprimido, mas não apresentou uma diferença significativa em relação ao controle. Pouco se sabe sobre a função do SigC, mas estudos relacionados a expressão, indicam que o gene é transcrito durante a fase de crescimento exponencial em *M. tuberculosis* (Manganelli et al., 1999). Para investigar se a deleção do gene *sigC* afeta a sobrevivência de *M. tuberculosis* em macrófagos, culturas do meio da fase de crescimento exponencial foram coletadas para infectar células do macrófago em camundongo. Os resultados indicaram que o fator sigma alternativo SigC exerce um papel central na letalidade de *M.tuberculosis* em camundongos (Sun et al., 2004).

Revelaram que o mutante *sigC*, faz parte da classe de mutantes em *M. tuberculosis* que persiste no tecido, mas são atenuados na capacidade para obter imunopatologia letal, o que implica que estes mutantes crescem e persistem nos pulmões dos camundongos, mas não causam mortalidade tão rapidamente como o tipo selvagem, e assim, a progressão histopatológica da doença é atrasada. Sendo assim, propuseram que mutantes imunopatológicos, como mutantes SigC, podem ser atraentes candidatos de vacina viva

atenuada, pois apresentam duas importantes características de um bom candidato vacinal: segurança e eficácia. O fenótipo imunopatológico tem sido observado também em mutantes para *sigE* e *sigH* em *M. tuberculosis* (Ando et al., 2003; Kaushal et al., 2002). Na fase estacionária *sigC* demonstrou regular genes envolvidos na sobrevivência de *M. tuberculosis* em macrófagos, *in vitro* (Yuan et al., 1998).

Como estamos simulando a condição de acidez no macrófago, é provável que o *sigC* tenha sido induzido para garantir a sobrevivência neste ambiente, mostrando que exerce uma função protetora, provavelmente, desde o início da fase exponencial, como Castro (Dissertação, 2009) observou em 15 min de exposição ao estresse oxidativo (um dos principais estresse no fagossomo), onde a expressão aumentou. Porém, observou que após 60 min a expressão de *sigC* foi equivalente a condição controle, devido ao crescimento que foi retomado e após 180 min, observou que *sigC* apresentou uma expressão reduzida. Contudo, observou que após 270 min, *sigC* apresentou acentuada expressão, o que levou o autor concluir que tenha ocorrido pela transição da cultura presente em condição normal, para a fase estacionária, enquanto a cultura do estresse permaneceu em crescimento exponencial e metabolicamente mais ativa. Tudo indica que na fase estacionária sua expressão seja realmente mais elevada.

No trabalho de Domingueti (Dissertação, 2011), através do emprego de plasmídeo suicida, foi possível avaliar a diferença no crescimento e viabilidade na linhagem mutante para *sigC* e selvagem (*C.pseudotuberculosis* 1002) frente a diferentes estresses e diferentes tempos, no início da fase exponencial ($DO_{600nm}=0,2$). Os resultados mostraram que a linhagem mutante foi mais sensível aos estresses oxidativo, osmótico e térmico que a selvagem indicando que SigC, possivelmente, desempenha importante papel na sobrevivência da bactéria no ambiente hostil. Contudo, após analisar as diferenças no crescimento relativo e na viabilidade relativa entre os tempos sucessivos de exposição das linhagens 1002 selvagem e mutante para o fator sigma C ao estresse ácido, utilizando-se o Teste T de Student, foi observado que esta condição de estresse afetou na mesma proporção a linhagem mutante e a selvagem e não afetou a viabilidade destas linhagens. Para esta autora, o fator sigma C não deve desempenhar um papel essencial para a resposta da bactéria ao estresse ácido.

Em nossos estudos, observamos que o *sigC* no estresse osmótico e térmico foi reprimido e não foi significativa a diferença em relação ao controle. Possivelmente, nestas condições o regulon de *sigC* não seja essencial ou até mesmo necessite ser reprimido para que outros genes possam atuar. E a expressão aumentada e significativa deste gene na condição de

estresse ácido, nos leva a supor que este gene tenha alguma contribuição para a resistência da bactéria neste ambiente.

Em *M. tuberculosis*, o mutante para *sigC* foi analisado quanto ao papel no seu regulon no início fase exponencial, início da estacionária e final da estacionária e o resultado mostrou redução na expressão de genes envolvidos em biossíntese de cofatores, componentes da parede celular, ácidos graxos e fosfolípidos, metabolismo de energia, conformação de proteína, síntese de proteína. Genes metabólicos são mais reprimidos durante o início da fase exponencial do que nas outras fases testadas (Sun et al., 2004). Observaram que genes associados à virulência foram reprimidos na ausência de *sigC*. Desvendar o regulon deste gene, em *C.pseudotuberculosis* 1002, seria essencial para descobrirmos quais genes estariam influenciando na sobrevivência e crescimento desta bactéria, principalmente em macrófagos e modelo animal.

O gene *sigD* foi considerado significativo no estresse osmótico e térmico, onde se mostrou induzido (tabela 16 e figura 24). No estresse ácido foi reprimido, mas não foi considerada uma diferença significativa em relação ao controle. Graham e Clark-Curtiss (1999) verificaram aumento significativo na expressão de *sigD* em fase tardia de crescimento comparada a fase inicial exponencial.

Um estudo envolvendo mutação no gene *sigD* em *M. tuberculosis* (Calamita et al., 2005) revelou que o mutante produz um fenótipo similar, mas não idêntico, aos mutantes imunopatológicos. Isto porque, o grau de atenuação não foi tão pronunciado como nos mutantes para *sigC*, *sigE* e *sigH* e também pela incapacidade de identificar uma diferença histopatológicas significativa nos tecidos de camundongos entre mutantes e linhagem selvagem. Contudo, foi possível observar atraso no tempo de morte dos camundongos, capacidade de replicação e persistência nos pulmões e exibiu a mesma taxa de crescimento da linhagem selvagem nos macrófagos. Pelas análises concluíram que o sigma D é dispensável ao crescimento *in vitro*, crescimento em macrófagos e crescimento e sobrevivência em camundongos, mas é requerido para letalidade do tipo selvagem neste animal.

Além disso, a perda de *sigD* está associado a redução na secreção de TNF- α a partir de macrófagos em cultura de tecido e é provável que seja a causa do atraso no tempo de morte. Os resultados também mostraram que SigD regula um conjunto de genes envolvidos com a síntese de proteínas ribossomais cuja expressão é, normalmente, induzida durante a fase estacionária, *in vitro* (Calamita et al., 2005).

O gene *sigE* codificante do suposto fator *sigE*, apresentou uma expressão de duas vezes a mais que o controle na condição de acidez (tabela 16 e figura 24). Nos estresses osmótico e choque térmico foi reduzido, mas não foi considerada uma diferença significativa em relação ao controle. Membros da subfamília do $\sigma 70$ com função extracitoplasmática, regulam a função relacionada à percepção e resposta às mudanças no periplasma e ambiente extracitoplasmático. O primeiro fator sigma ECF identificado em *E.coli* foi o σE , o qual foi reconhecido como segundo fator sigma de resposta ao choque térmico no organismo (Wang e Kaguni, 1989). Em *Haemophilus influenzae* a expressão de *rpoE*, aumentou em 102 vezes dentro de macrófagos e a sobrevivência de mutantes foi reduzida a níveis relativos ao tipo selvagem (Craig, Nobbs, High, 2002). Mutantes para *rpoE* em *Vibrio cholerae* são virulentos atenuados, exibindo capacidade reduzida em colonizar intestino de camundongos (Kovacikova e Skorupski, 2002).

Segundo Pacheco (dados não publicados), ao aplicar diferentes estresses na cultura de *C.pseudotuberculosis* 1002 em fases de crescimento diferentes e em culturas mais jovens ($DO_{600nm}=0.5-0.6$) foi possível perceber que a linhagem mutante 1002($\Delta sigE$) apresentou maior sensibilidade que a linhagem selvagem a agentes geradores de estresse de superfície celular: etanol e SDS.

No trabalho com mutante para *sigE* em *M. tuberculosis* H37Rv, Manganelli et al (2001) demonstraram que o mutante foi mais sensível a vários estresses ambientais, como choque térmico, SDS, agentes oxidativos, mas não para acidez. Porém, em estudos contraditórios, Wu et al (1997) relataram que o mutante *sigE* em *M. smegmatis* é mais sensível a acidez, peróxido de hidrogênio, SDS e choque térmico. Ainda no trabalho com *M. tuberculosis* H37Rv Manganelli et al (2001) verificaram que o *sigE* é necessário para resposta ao estresse, bem como à capacidade de crescer e sobreviver dentro dos macrófagos.

Este fator também influencia o nível de *sigB* na célula, mas a expressão de *sigB* não é completamente influenciada pelo *sigE*, uma vez que o nível de *sigB* reduziu na linhagem mutante, enquanto que o nível de mRNA de outros sigmas não foram afetados, porém, logo após, *sigB* foi induzido no choque térmico, não permanecendo reprimido. Um total de 38 genes mostrou ser influenciado pelo *sigE*, incluindo regulador transcricional, enzimas envolvidas na degradação de ácido graxo e proteínas clássicas de choque térmico (Manganelli et al., 2001).

Em *C.pseudotuberculosis* 1002, na condição de acidez, *sigE* foi induzido e significativo, bem como o *sigB*, enquanto que, nas condições onde o nível de mRNA de *sigE* não foi considerado significativo, *sigB* também não foi. Pode haver alguma influência entre os

dois, nesta fase de crescimento, porém estudos mais específicos serão necessários para se chegar a uma conclusão mais precisa.

O gene *sigH*, codificante do fator sigma H, foi induzido e considerado significativo em todas as condições e na acidez foi altamente expresso (tabela 16 e figura 24). Em *M.tuberculosis* o mutante para *sigH* exibiu alta susceptibilidade para estresse oxidativo e choque térmico *in vitro*, sugerindo importante papel do *sigH* na rede regulatória envolvendo resposta ao estresse (Raman et al., 2001). No trabalho de Castro (Dissertação, 2009) foi demonstrado que em 15 min no estresse oxidativo proveniente de H₂O₂, a expressão deste gene aumentou e foi ainda maior no tempo de 60 min, sendo que nos tempos de 180 min e 270 min, a expressão se apresentou similar ao controle.

Em *Corynebacterium glutamicum*, este gene é responsável pela expressão de Clp, uma protease envolvida na degradação de proteínas não funcionais na célula, normalmente originadas pelo choque térmico (Engels et al., 2004). Uma mutação no gene *sigH*, interrompeu a transcrição de *sigM*, e a detecção do promotor de *sigH* acima de *sigM* implica que este fator está sob controle transcricional direto de *sigH* (Nakunst et al., 2007).

Em *C.pseudotuberculosis* 1002, a condição de estresse onde *sigH* aumentou, *sigM* também aumentou, mas nesta fase de crescimento os transcritos de *sigM* não apresentam uma indução tão relevante, se compararmos com a expressão dos outros genes codificantes de sigmas. O gene *sigH* está localizado (fita direta) abaixo do anti-sigma Cp1002_0531 (RshA), o qual foi induzido e significativo em todas as condições. No estresse ácido apresentou uma expressão sete vezes maior em relação ao controle (anexo 8). Quando mudanças no ambiente são percebidas pelo fator anti-sigma, o complexo sigma-anti-sigma se desfaz, tornando o sigma ativo, resultando na ligação do core de RNA polimerase e ativando a transcrição dos genes regulados por ele (Hughes e Mathee, 1998).

Em *M. tuberculosis*, a fosforilação do RshA, pela proteína quinase serina/treonina (*pknB*) reduziu a interação com SigH, *in vitro*, confirmando a importante participação na adaptação ao estresse oxidativo nesta bactéria. A atividade de SigH é essencial para virulência em *M. tuberculosis*, e esta regulação mediada pela quinase pode ser crítica para patogenicidade da tuberculose. Em *C.pseudotuberculosis* 1002, o gene *pknB* foi induzido e significativo em todas as condições, principalmente na acidez (anexo 8). O que se pode inferir que este gene tenha importante participação na regulação de *sigH*, pois na condição onde o gene se mostrou mais induzido, *sigH* também se mostrou mais induzido juntamente com o anti-sigma, que também tem a função de sensor na alteração do ambiente externo. Os genes *pknA* e *pknB* são

co-transcritos e participam na regulação da morfologia celular e síntese da parede celular (Kang et al., 2005).

No trabalho de Souza (Dissertação, 2011), foi construído uma linhagem mutante de *C.pseudotuberculosis* 1002 deficiente para o gene *sigH* e submetida, juntamente com a linhagem selvagem, a diferentes condições de estresse e não foram observadas diferenças significativas na susceptibilidade da linhagem mutante para o fator σ^H aos estresses ácido, térmico e oxidativo em relação à linhagem 1002 selvagem. Contudo, foi verificado que esta linhagem é significativamente mais susceptível ao estresse osmótico do que a linhagem 1002 selvagem, indicando que o fator σ^H deve desempenhar um papel importante para a resposta da bactéria a esta condição de estresse.

O gene *sigM* apresentou uma expressão significativa no estresse osmótico e acidez com *fold-change* de aproximadamente 5 nas duas condições (tabela 16 e figura 24). Em choque térmico não foi considerada uma diferença significativa, mas foi induzido em relação ao controle. Possivelmente, este gene não é preferencialmente expresso no início da fase exponencial. Análises de microarranjo em *C.glutamicum* na fase exponencial ($DO_{610nm} = 0,2$ a $0,3$) revelou que o rompimento de *sigM*, não afetou o nível de transcrição dos genes induzidos por choque térmico, o que implica que este fator sigma não está envolvido na regulação da expressão de genes em resposta ao choque térmico (Ehira et al., 2009).

Em *M. tuberculosis*, por exemplo, a expressão ocorreu na fase estacionária e somente diante do estresse térmico (Agarwal et al., 2007). Castro (Dissertação, 2009) em seu trabalho com *C.pseudotuberculosis* 1002 observou um aumento de expressão deste gene somente após 270 min de aplicação do estresse oxidativo. Em *C. glutamicum*, experimentos revelaram que uma deleção do *sigM* causou redução no número de células viáveis diante da aplicação do estresse por choque térmico, frio e dissulfeto (subtipo do estresse oxidativo) na cultura em fase de crescimento exponencial $DO_{600nm}=0,7$. Além disso, os experimentos com PCR em tempo real demonstraram que a transcrição de *sigM* aumentou significativamente depois da aplicação destes estresses. Estes resultados sugerem que este fator está envolvido na resposta ao estresse (Nakunst et al., 2007), contudo, com forte indício que seja em fase mais tardia de crescimento.

O gene *sigK* foi considerado significativo apenas no estresse térmico, mas induzido em todas as condições em relação ao controle (tabela 16 e figura 24). Similar ao gene *sigM*, não apresentou número de transcritos tão consideráveis como nos outros genes codificantes de sigma. Castro (Dissertação, 2009) somente detectou forte indução deste gene em relação ao controle, após 270 min de exposição ao H_2O_2 . Como parece ocorrer em *sigM*, este resultado

pode ser indício que este fator esteja envolvido na resposta tardia ao estresse e é o único sigma que foi predito em uma das ilhas de *C. pseudotuberculosis* (tabela 17). Para todos os genes codificantes dos fatores sigmas, faz-se necessário desvendar o regulon e revelar também os prováveis reguladores destes genes, ou seja, revelar toda a rede de interação dos genes codificantes de fatores sigmas para melhor entendimento do processo de infecção e modulação da resposta na célula.

6.7. Genes das Ilhas de Patogenicidade preditas no genoma de *C.pseudotuberculosis* 1002

Os genes que contribuem para a virulência são induzidos somente em contato com hospedeiro ou invadindo o hospedeiro. A expressão de tais genes *in vivo* é dependente do patógeno apresentar a capacidade para perceber seu ambiente imediatamente, que seja suficiente para identificar o contato no hospedeiro. Estes genes em sua grande maioria estão presentes em ilhas de patogenicidade (Webb e Kahler, 2008).

A maioria dos genes mais caracterizados de *C.pseudotuberculosis* 1002 estão presentes nas ilhas de patogenicidade pela predição de Soares (Dissertação, 2009), incluindo a *pld*, operon *fagABC* e gene *fagD*. Na análise do transcriptoma, pode-se observar na tabela 17, que a maioria dos genes presentes nas ilhas, não foram considerados diferencialmente expressos, em pelo menos uma condição analisada. Na figura 25 verificamos os genes diferencialmente expressos em cada condição testada. Todos os genes presentes nas ilhas sugerem forte participação na virulência desta bactéria. De 57 genes diferencialmente expressos nas ilhas em pelo menos uma condição, 26 genes apresentaram um RPKM acima de 100, considerando pelo menos uma condição. O RPKM está baixo para a maioria dos genes e é bem provável, que entre os fatores envolvidos na regulação destes genes (em sua maioria desconhecidos), a densidade celular esteja atuando. Descrevemos alguns genes abaixo.

O gene *cat*, codificante da catase foi altamente expresso na condição de acidez pH= 5. O gene apresentou um valor de *fold-change* 4 vezes maior que o controle, com uma cobertura alta em todas as situações. Os mecanismos usados pela bactéria para resistir aos prejuízos causados por espécie reativa de oxigênio são pontos centrais para a sobrevivência dentro do macrófago. Estes mecanismos são mediados, provavelmente, pela expressão de catalase/peroxidase. O papel da catalase é reduzir a concentração de peróxido de hidrogênio

(H₂O₂) na célula. A molécula de O₂ tem sempre o mesmo número de elétrons, mas a sua reatividade aumenta quando aceita um, dois, ou três elétrons para formar, respectivamente, um radical superóxido (O₂⁻), peróxido de hidrogênio (H₂O₂) e radicais hidroxilas (OH⁻) (Ratledge e Dover, 2000).

O radical livre da espécie química de oxigênio pode prejudicar as células pela oxidação da membrana, alvos periplasmáticos, enzimas, proteínas, DNA e grupos sulfato ferroso. Outras espécies reativas de oxigênio não radical, incluindo oxigênio singlete, ozônio, peroxidação lipídica, ácido hipocloroso podem também prejudicar a célula (Wesche et al., 2009). As moléculas ativas do oxigênio são produzidas como subproduto inevitável do metabolismo aeróbico normal, e são acentuadas pela exposição a certos ambientes (Farr e Kogoma, 1991). Em ambiente aquoso neutro, a dismutação espontânea de O₂⁻ produz H₂O₂. Sob condições mais ácidas, a molécula O₂⁻, pode ser protonada para formar HOO⁻ (radical hidroperoxil) (Farr e Kogoma, 1991), que ao ser reduzido, levará a produção de H₂O₂. Este radical (O₂⁻) é muito mais reativo por causa da carga negativa a ser neutralizada. Quando o pH da célula diminui, a razão de HOO⁻ para O₂⁻ aumenta, conseqüentemente, para sua sobrevivência, a bactéria tem que aumentar a produção de catalase para combater a grande quantidade de peróxido produzido.

Foi demonstrado por Imlay e Fridovich (1991) que a principal fonte de O₂⁻ em *E.coli*, é a auto-oxidação de NADH desidrogenase, succinato desidrogenase, e D-lactato desidrogenase. Uma enzima citosólica que acredita ser capaz de gerar uma quantidade significativa de O₂⁻ é a glutathiona redutase, que usa NADH como fonte de elétron. Em *C.pseudotuberculosis* 1002 esta referida enzima foi induzida em todas as condições, porém com uma razão abaixo de 2 e por isso, somente será visualizado o valor de expressão, na tabela que contém os genes comuns, independentes do *fold-change*, no anexo 8. Na condição de acidez esta enzima apresentou valor um pouco maior, aumentando ainda mais a probabilidade de gerar radical superóxido neste ambiente.

O tratamento térmico pode inativar a catalase e superóxido dismutase em vários microrganismos, incluindo *L. monocytogenes* e *S. aureus*, o que em conseqüência, leva ao aumento de radical hidroxila, prejudicial à célula (Andrews e Martin, 1979). Como neste trabalho simulou-se o ambiente do macrófago, este resultado demonstra forte indício da produção de catalase no início da fase exponencial, para escape do sistema de defesa do hospedeiro nesta célula, onde ocorre grande produção do peróxido de hidrogênio, visando eliminar a célula bacteriana.

Pode-se sugerir que o regulador LysR, cujo gene foi induzido na condição de acidez (tabela 12) e está envolvido na regulação de resposta oxidativa possa ser o regulador positivo da catalase, uma vez que para seu homólogo, OxyR, já existe relato do envolvimento na regulação positiva da catalase, onde o mutante *oxyR* mostrou ausência da enzima em fase estacionária em *Caulobacter crescentus* (Italiani et al., 2011). Em *C. pseudotuberculosis* 1002, o gene *lysR* na condição de estresse osmótico foi reprimido e não foi considerado significativo, assim como o gene *cat* nesta condição. E no estresse térmico, o gene *lysR* foi considerado significativo e induzido, como o gene *cat* com valor de expressão do estresse pouco acima do controle, tanto para *lysR* quanto para *cat*, apresentando um *fold-change* pouco acima de 1.

Vale ressaltar que para cada espécie a regulação pode alterar, assim, faz-se necessário conhecer toda a rede de regulação envolvendo a catalase na espécie *C.pseudotuberculosis* 1002 e verificar esta possível participação de LysR na regulação deste gene. Em *E.coli* e *Salmonella enterica*, por exemplo, genes ativados pelo peróxido de hidrogênio, incluindo catalase, apresentam o gene OxyR como regulador positivo, porém em *Neisseria gonorrhoeae*, este gene foi inativado e a expressão do gene codificante de catalase aumentou 9 vezes e o mutante se mostrou mais resistente aos danos causados pelo peróxido de hidrogênio que a linhagem selvagem, demonstrando que OxyR atua como repressor nesta espécie (Tseng et al., 2003).

Uma observação interessante é que no genoma de *C.pseudotuberculosis* 1002, o gene *sigC* está localizado na fita reversa há uma distância de apenas 47 aminoácidos do gene codificante da catalase, que está na fita direta, e está igualmente envolvido na proteção contra o estresse oxidativo. Então poderia-se sugerir que *sigC* esteja envolvido na regulação de catalase. Porém, ao analisar os valores de expressão entre *sigC* (tabela 16) e *cat* (tabela 17) no estresse térmico, houve alternância de expressão entre eles, onde o gene *sigC* foi reprimido enquanto *cat* foi induzido. Se existir alguma participação de SigC na regulação de *cat*, pode-se ainda sugerir que seja devido a uma dependência da interação entre LysR e SigC na determinação da ativação do gene *cat*, uma vez que os fatores sigmas fornecem um local de contato para algumas proteínas ativadoras e esta interação possivelmente aumenta a habilidade da ligação de RNA polimerase ao promotor (Helmann, 2005) e assim, poderia ser uma situação específica da acidez. São hipóteses que devem ser analisadas. Contudo, pelos resultados apresentados é bem provável que a expressão do gene *cat* seja mesmo regulada por LysR.

Muitas proteínas hipotéticas apresentaram um comportamento interessante e diferenciado em cada condição. A maioria não foi considerada diferencialmente expressa em

todas as condições, apesar da diferença de expressão nos estresses em relação ao controle. Apresentaram-se com bastante plasticidade de expressão entre as condições. Estas proteínas precisam ser investigadas, pois podem se tornar importantes alvos vacinais. Como exemplo citamos: Cp1002_0048, Cp1002_0174, Cp1002_0176, Cp1002_0177, Cp1002_0991, Cp1002_1454.

Em relação ao cluster gênico da urease, somente o *ureG* em *C.pseudotuberculosis* 1002, foi diferencialmente expresso nas três condições e induzido. Esta proteína está envolvida na catalização da hidrólise de urea para produção de amônia e carbamato, e este último composto é espontaneamente degradado para produzir outra molécula de amônia e ácido carbônico. Em solução, tanto o ácido carbônico liberado quanto duas moléculas de amônia estão em equilíbrio com suas formas desprotonada e protonada, respectivamente, tanto que o efeito desta reação é o aumento de pH (Mobley Island, Hausinger, 1995). Em *Helicobacter pylori* o cluster consiste de genes estruturais *ureAB*, que codificam subunidades da enzima e genes acessórios *ureIEFGH*, que são requeridos para a ativação da urease (Muckle e Gyles, 1982)(Cussac, Ferrero, Labigne, 1992). Em *C.pseudotuberculosis* 1002, *ureABC* supostamente indicam genes estruturais, e *ureEFGD* genes acessórios.

O gene *ureA*, significativo ($p\text{-value} < 0,001$) apenas no estresse ácido, foi também o único que apresentou baixa cobertura dos transcritos, apresentando um RPKM baixo. O gene *ureD* foi induzido no ambiente osmótico e térmico. O gene *ureE* foi induzido no estresse térmico e *ureF* no estresse osmótico. Em nenhuma condição o sistema foi induzido ou reprimido por completo. Estudos com espécies bacterianas têm mostrado que a expressão de urease, que é uma enzima ligante ao níquel, é ativada por H-NS (proteína associada a nucleóide-), UreR (gene regulatório), NikR (regulador de resposta ao níquel), e ArsR (repressor ArsR) (Coker, Bakare, Mobley 2000).

Em *C. pseudotuberculosis* 1002 o gene anotado como *arsR*, que é próximo ao cluster não foi considerado significativo em nenhuma condição, bem como os outros genes codificantes da proteína estrutural (*ureB* e *ureC*) também. Para Kirk e Sumner (1934), a urease (que foi isolada de feijão) pode ser usada como antígeno capaz de estimular uma forte resposta imunológica. Soro de coelhos imunizados com a enzima purificada foi capaz de inibir a atividade da urease em uma forma titulável.

O sistema urease é o processo básico utilizado para manter o pH do periplasma e citoplasma, próximo a neutralidade, quando *H.pylori* é exposta às condições ácidas. Em pH 5.0, *ureA*, *ureB* e *ureG* foram induzidos, mesmo quando não foi adicionado uréia ao meio

(Shao et al., 2008). A indução do gene *ureA* e *ureB* não foi surpresa, pois são co-transcritos a partir do mesmo promotor e ambos são essenciais para a atividade da urease (McGee et al., 1999). Em *C.pseudotuberculosis* 1002, esta situação não ocorreu. No meio ácido, o gene *ureA* foi induzido enquanto *ureB* foi reprimido, levando a supor uma plasticidade de expressão e indicando promotor diferente ou um regulador atuando diferentemente em cada um.

Em *H.pylori* linhagem 26695, Shao et al (2008) estudaram o perfil proteômico, onde a bactéria foi exposta a diferentes níveis de pHs na ausência de urease, usando gel bidimensional, buscando identificar outras enzimas relacionadas a acidez. Através da análise por MALDI-TOF-TOF-MS, conseguiram identificar 36 proteínas, cuja função incluía produção de amônia, chaperonas, metabolismo de energia, regulador de resposta, envelope celular e algumas proteínas com função desconhecida.

É provável que em *C.pseudotuberculosis* 1002, outros mecanismos também estejam ativos nesta condição, pois o sistema de urease, além de não se mostrar completo nos transcritos, em nenhuma condição nesta fase de crescimento, apareceu com um valor de expressão tão significativo, mas a bactéria conseguiu sobreviver cerca de 70%, em cada condição.

O gene *lcos*, codifica uma Two-component sensor protein, sensor histidine kinase, que pertence a família de histidina quinase, ao qual o domínio *envz* de *E.coli* também faz parte. A proteína EnvZ é uma histidina quinase que é autofosforilada por um grupo fosfato originado do ATP em um resíduo de histidina conservado. Subsequentemente transferem o grupo fosforil para o resíduo aceptor aspartato de proteína reguladora de resposta (Paithoonrangsarid et al., 2004). A proteína sensora de dois componentes presente em *C.pseudotuberculosis* 1002 apresenta o domínio de dimerização HisKA e o C-terminal de ligação ao ATP ou de fosfotransferência, o HATPase_C.

No genoma de *C.pseudotuberculosis* 1002, abaixo do gene *lcos*, está localizado o gene *phoB*, que codifica supostamente Two-component regulatory protein. Este gene apresenta o domínio membro da superfamília *trans_reg_C*, o qual *ompR*, faz parte. A fosforilação então ativa o domínio efetor do regulador de resposta, que arma a resposta celular. O sistema de dois componentes está envolvido na percepção e transdução de sinais em diferentes condições de ambientes, especialmente ambiente hiperosmótico (Stock, Robinson, Goudreau 2000).

Em *C.pseudotuberculosis* 1002, o sistema *lcos-phoB* foi considerado diferencialmente expresso apenas na condição de estresse osmótico e foi induzido, indicando que o sistema é acionado logo no início da fase exponencial. No momento que entra em contato com o

ambiente osmótico no hospedeiro, é muito provável que a bactéria já acione uma resposta específica para conseguir sobreviver neste ambiente.

O gene *mnhG* apresentou um RPKM baixo e considerado induzido apenas na condição de acidez. Apesar de que, na condição de estresse térmico, o valor está maior que na acidez, mas não foi considerado significativo. Possivelmente se trata de um viés do programa DEGseq. O gene codifica a suposta subunidade G do sistema antiporter Na^+/H^+ , onde o operon em *C.pseudotuberculosis* 1002 é constituído por seis genes (subunidades BCDEFG). Em *Staphylococcus aureus* são sete ORFs constituindo o operon (*mnhA* até *mnhG*) (Hiramatsu et al., 1998). Os íons mais comuns na célula são Na^+ e H^+ e desempenham papel primordial na fisiologia da célula: ambos são os mais importantes bioenergéticos celulares e importantes por manter o equilíbrio na concentração de prótons dentro da célula, que é crítica pra o funcionamento celular e suas proteínas.

Na verdade, quando a concentração destes íons torna-se muito alta ou muito baixa, eles se transformam em fatores de força potente para todas as células. Assim, cada célula tem um mecanismo homeostático eficiente para estes íons. As proteínas que exercem este mecanismo são antiporter Na^+/H^+ (Padan et al., 2001). O antiporter tem várias funções na célula bacteriana: i) estabelecer o potencial eletroquímico de Na^+ através da membrana citoplasmática, ii) troca de Na^+ (ou Li^+), que são tóxicos se acumulados em alta concentração na célula, para H^+ , iii) regulação do pH intracelular e iv) regulação do volume celular (Hiramatsu et al., 1998). Como em *C.pseudotuberculosis* 1002 foi induzido na condição de acidez, pode-se supor que no macrófago este gene seja logo ativado para a sobrevivência neste ambiente, com alta concentração de íons H^+ .

O gene Cp1002_1467 codifica o suposto receptor hemina da superfície celular. Esta proteína apresenta o domínio encontrado em HtaA, que é uma proteína secretada responsável pela aquisição do ferro e transporte. Este gene foi considerado reprimido em todas as condições, mas só foi considerado significativo na condição do estresse osmótico. *Corynebacterium* pode adquirir ferro a partir do grupo hemina e hemoglobina durante o crescimento, em meio onde não existe ferro livre disponível (Schmitt, 1997). Mutantes de *C.ulcerans* e *C.diphtheriae* defeituosos na utilização do ferro a partir de hemina foram incapazes de utilizar este grupo como fonte de ferro, no meio com baixa de ferro livre, mas na presença de hemina ou hemoglobina. (Drazek, Hammack, Schmitt, 2000). Ferro é essencial para as bactérias e é requerido por muitas bactérias patogênicas para sobreviver no hospedeiro

e provocar doenças. Grande parte do ferro em mamíferos é sequestrado, sendo a maioria encontrado intracelularmente associado com hemina na proteína hemoglobina.

A hemina liberada pela lise de eritrócitos é rapidamente ligada por proteínas do soro, como albumina, enquanto o ferro extracelular é logo associado às proteínas ligantes ao ferro por alta afinidade, como transferrina (Weinberg, 1978). Muitas bactérias patogênicas em condições limitantes do ferro no hospedeiro utilizam desta situação como sinal para induzir a expressão de vários determinantes de virulência (Litwin e Calderwood, 1993). Mecanismos independentes de sideróforos para aquisição de ferro no hospedeiro inclui sistemas envolvidos na extração de ferro a partir de transferrina e/ou lactoferrina (Mietzner e Morse, 1994) ou adquire a partir de hemina e de proteínas contendo hemina através de receptores específicos. Hemina é uma forma oxidada da heme que é encontrado no meio extracelular (Francis et al., 1985),

Apesar de tanto bactéria gram-negativa quanto gram positiva empregarem o transportador do tipo ABC e proteína ligante associada ao substrato para mover a hemina através da membrana citoplasmática, estes organismos utilizam componentes diferentes para ligar hemina ou hemoglobina na superfície celular. Gram-negativas ligam estes grupos em grandes proteínas da membrana externa e gram-positivas, ligam hemina ou hemoglobina em proteínas ancoradas no envelope celular. Estes receptores ou são ligados covalentemente no peptidoglicano da parede via sortases, ou são presos na membrana citoplasmática através da região C-terminal de uma transmembrana (Allen e Schmitt, 2009).

Em *C.pseudotuberculosis* 1002, o gene codificante de uma proteína com suposta função de utilizar o ferro a partir de hemina foi reprimido em todas as condições. No meio BHI, pode haver uma pequena quantidade de ferro, oriunda da composição contendo infusão de bife, coração, de cérebro de bezerro (possivelmente contendo sangue). Nesta ocasião, pode-se supor que o ferro esteve presente em uma quantidade que não foi suficiente para ativar o gene, apesar de normalmente os captadores de ferro serem ativos exatamente pela baixa concentração de ferro ou ainda, este gene é também dependente de densidade celular. Estudos mais específicos devem ser realizados para desvendar a regulação deste gene.

Intrigante pois o gene *phuC*, que também é responsável pela captação de ferro a partir do grupo heme ou associado a sideróforos, foi induzido em todas as condições, como se pode observar na tabela 8 do core estimulon. Possivelmente um dos fatores de regulação deste receptor hemina seja mesmo dependência da densidade celular.

6.8 Produtos gênicos confirmados, *in silico* ou experimental, por outros trabalhos do grupo LGCM, envolvendo *C.pseudotuberculosis* 1002

No Laboratório de Genética Celular e Molecular (LGCM), são várias as linhas de pesquisa envolvendo *C.pseudotuberculosis* de diferentes linhagens. Agrupando os trabalhos realizados com a linhagem 1002, até o momento, incluindo predição *in silico* de localização sub-celular de proteínas através do SurfG+, imunoproteômica, exoproteoma com duas tecnologias diferentes e phage display, observamos (anexo 15) como se comportam estes genes no início da fase exponencial diante dos estresses aplicados.

Atentamos em averiguar os genes com *fold-change* igual ou maior a 2 nas três condições. Os gene *sdrC* e o gene Cp1002_0164, apresentaram um *fold-change* no corte estabelecido, com valor alto no estresse osmótico e acidez, e no térmico, apesar de ter apresentado um *fold-change* de 2, não foi considerado significativo (anexo 15). Estes genes foram preditos apenas no trabalho *in-silico*, através do programa SurfG+ (anexo 15). O gene Cp1002_0893 codificante de suposta proteína secretada foi significativo em todas as condições com *fold-change* acima de dois e merece atenção em estudos posteriores relacionados ao tratamento e controle da doença. Este gene foi detectado por SurfG+ e Phage Display (anexo 15). Os genes de ocorrência em todos os trabalhos, ao mesmo tempo, foram os que apresentaram menor expressão nesta fase de crescimento da bactéria.

O gene *sdrC* codifica uma proteína contendo repetição de Serina-aspartato. Em *S.aureus* esta subfamília está associada com a parede celular, com potencial função de adesina (Josefsson et al., 1998). A organização dos domínios se assemelha às proteínas de superfície ClfA e ClfB em Gram-positivas, que são fatores de virulência que favorecem a ligação da bactéria em fibrinogênios (Mcdevitt et al., 1994). Uma longa sequência sinal (S) é seguida por uma região (A) de aproximadamente 500 resíduos que abrange uma sequência única. Uma região R, que é composta da repetição com dipeptídeo serina-aspartato (SD), liga na região A para ancorar na parede celular a região W. ClfA e ClfB são ancoradas na parede celular via motivo LPXTG presente no domínio W (Schneewind et al., 1995). A extremidade C-terminal inclui uma região transmembrana (M) e uma cauda curta citoplasmática composta de resíduos com carga positiva (Josefsson et al., 1998)

A região A destas proteínas SdrC são relacionadas o que comprovam origem comum, mas elas são suficientemente variadas sugerindo que podem se ligar em diferentes ligantes e tem diferentes funções. A região A de ClfA e ClfB se liga no fibrinogênio, mas a região A de

SdrC interage com proteínas do hospedeiro ainda não identificadas (Josefsson et al., 1998). Assim, este gene torna-se alvo de estudos para testes de possíveis candidatos vacinais.

Interessante que o outro gene selecionado, Cp1002_0164, codificante de uma suposta proteína secretada com motivo LPXTG foi induzido exatamente nas mesmas condições que o *sdrC* que está *downstream* ao gene Cp1002_0164 na fita direta, em uma distância de apenas 203 aminoácidos. É possível que sejam transcritos juntos e pela atividade exercida por proteínas com motivo LPXTG, relatada no item 6.2, este gene torna-se também importante alvo vacinal.

6.9. Ilustração de um possível modelo de regulação entre alguns genes diferencialmente expressos

A partir dos dados obtidos foi possível relacionar alguns genes diferencialmente expressos em cada condição analisada e no core estimulon, conforme dados da literatura ou suposições baseadas na localização dos genes no genoma de *C.pseudotuberculosis* 1002 e pelo envolvimento ao mesmo processo biológico, como *GntR*, *nanK* e *opp*. Na figura 27 propomos um modelo de regulação e representamos alguns genes que se destacaram pelo valor de RPKM apresentado em cada condição.

7. CONCLUSÕES

O presente trabalho é pioneiro no Brasil e amplia o conhecimento sobre os recursos utilizados pela bactéria no processo de infecção e abre um leque para novas investigações biológicas, visando melhor entender sua regulação e patogenia. Através do sequenciamento do cDNA realizado pela plataforma SOLiD™, uma tecnologia de nova geração, foi possível identificar grande número de genes diferencialmente expressos, e verificar provável interação entre eles;

Todo o genoma de *C.pseudotuberculosis* 1002 foi representado conseguindo distinguir regiões transcricionalmente ativas em todas as condições testadas e observando diferenças na expressão de cada gene, verificando a influência dos estresses na expressão de cada um;

O estresse térmico, depois de estabelecido o corte pelo *fold-change*, foi o meio que menos representou genes, levando a sugerir que a maioria dos genes são, possivelmente, regulados, além de outros fatores, pela densidade celular. Porém, genes interessantes foram descobertos como o gene *ahpD*, que em outro microrganismo já demonstrou seu papel protetor. Genes codificantes de chaperonas também foram induzidos, como já era esperado;

O estresse osmótico e ácido apresentou quase a mesma quantidade de genes diferencialmente expressos, mostrando que logo no início a bactéria se prepara para se manter e sobreviver nestes ambientes prejudiciais para a célula;

No estresse osmótico, entre outros genes interessantes, o destaque foi para o gene *spaD* que, quando interrompido em *C. pseudotuberculosis* linhagem T1 virulenta, apresentou proteção de 60% contra LC em trabalho realizado por membro da nossa equipe;

O estresse ácido foi a condição que mais apresentou genes relacionados à virulência, como o *fagD*, *dps*, *msrB*, *lysR* que além de contribuírem com a informação a respeito do processo de infecção, podem se tornar fortes candidatos vacinais;

Vários genes parecem ser regulados pela densidade celular, e possivelmente a expressão seja maior em fase de crescimento mais tardia, mas ainda assim, nossos resultados revelaram vários genes que demonstram ser importantes para sobrevivência e virulência dentro de cada nicho testado, principalmente proteínas hipotéticas.

Acreditamos que os resultados advindos através deste projeto sejam primordiais para entendimento do desenvolvimento da infecção causada por *C. pseudotuberculosis* e que podem ser utilizados como alvos para tratamento e desenvolvimento de novas vacinas contra este importante microrganismo, em estudos futuros.

8. PERSPECTIVA

Com os dados obtidos pelo sequenciamento, pretendemos:

- Identificar os transcritos não codificantes que possivelmente estejam envolvidos na regulação,
- Aprimorar o processo de anotação,
- Fazer a predição da rede regulatória *in silico* e validar com estes dados experimentais
- Realizar réplicas biológicas (duas) para validação dos resultados com a mesma tecnologia utilizada neste trabalho

Já finalizada toda a parte experimental relacionada a linhagem 258, a próxima etapa visa:

- Identificar o perfil transcricional desta linhagem e comparar os dois microrganismos (1002 e 258) para detectar a preferência ao hospedeiro, ou maior resistência e sensibilidade em determinado meio.

9. REFERÊNCIAS BIBLIOGRÁFICAS

ADAMCZYK-POPLAWSKA MONIKA, MARKOWICZ SERGIUSZ, JAGUSZTYN-KRYNICKA ELZBIETA K., **Proteomics for development of vaccine**, *Journal of Proteomics* 2011

AGAFONOV, D. E.; KOLB, V. A.; SPIRIN, A. S. Ribosome-associated protein that inhibits translation at the aminoacyl-tRNA binding stage. **EMBO reports**, v. 2, p. 399-402, 2001.

AIRELLO, S.; MAYS, A.; OLIVEIRA, P.M.A. Manual Merck de Veterinária. 2001. Editora ROCA, São Paulo, 8ªed.

ALAMURI, P.; MAIER, R. J. Methionine sulphoxide reductase is an important antioxidant enzyme in the gastric pathogen *Helicobacter pylori*. **Molecular microbiology**, v. 53, p. 1397-406, 2004.

ALAMURI, P.; MAIER, R. J. Methionine Sulfoxide Reductase in *Helicobacter pylori*: Interaction with Methionine-Rich Proteins and Stress-Induced Expression. **J. Bacteriol.**, v. 188, n. 16, p. 5839-5850, August 15, 2006

ALEXANDER, M. **Biodegradation and Bioremediation**. 2 edition. Academic Press., 453. 1999

ALLEN, C. E.; SCHMITT, M. P. HtaA is an iron-regulated hemin binding protein involved in the utilization of heme iron in *Corynebacterium diphtheriae*. **Journal of bacteriology**, v. 191, p. 2638-48, 2009.

ALLOING, G.; DE PHILIP, P.; CLAVERY, J.-P. Three Highly Homologous Membrane-bound Lipoproteins Participate in Oligopeptide Transport by the Ami System of the Gram-positive *Streptococcus pneumoniae*. **Journal of molecular biology**, v. 241, n. 1, p. 44-58, 1994.

ALVES, F.S.F., OLANDER H. Uso de vacina toxóide no controle da LC em caprinos. **Veterinária Notícias**, p. 69-75. 1999

ALVES, F.S.F.; PINHEIRO, R.R.; PIRES, P.C. LC: patogenia-diagnóstico- controle. Artigo Técnico. Documento nº 27, Embrapa Sobral CE, **Ministério da Agricultura e do Abastecimento**. 1997.

ALVES, F.S.F., PINHEIRO, R.R. LC em caprinos e ovinos: recomendações e medidas profiláticas. **Agropecuária Catarinense**, v.13, n.1, p.12-14, 2000.

ANDO M, YOSHIMATSU T, KO C, CONVERSE PJ, BISHAI WR. Deletion of *Mycobacterium tuberculosis* sigma factor E results in delayed time to death with bacterial persistence in lungs of aerosol-infected mice. **Infect Immun** 71: 7170-7172. 2003

ANDREWS GP, MARTIN SE. Heat inactivation of catalase from *Staphylococcus aureus* MF-31. **Appl Environ Microbiol** 37(6):1180-5. 1979

APPLIED BIOSYSTEMS BY LIFE TECHNOLOGIES. Disponível em <<http://www.appliedbiosystems.com.br>>. Acessado em 26 de maio, 2011.

ARGOS, P.; HANEL, M.; WILSON, J. M.; KELLEY, W. N. A possible nucleotide-binding domain in the tertiary fold of phosphoribosyltransferases. **Journal of Biological Chemistry**, v. 258, n. 10, p. 6450-7, May 25, 1983 1983.

AYERS JL. Caseous lymphadenitis in goat and sheep: review of diagnosis, pathogenesis, and immunity. **JAVMA**, n. 171: pp. 1251-4. 1977

BABU, M. M.; PRIYA, M. L.; SELVAN, A. T.; MADERA, M.; GOUGH, J.; ARAVIND, L.; SANKARAN, K. A Database of Bacterial Lipoproteins (DOLOP) with Functional Assignments to Predicted Lipoproteins. **J. Bacteriol.**, v. 188, n. 8, p. 2761-2773, April 15, 2006 2006.

BAIRD, G. J.; FONTAINE, M. C. *Corynebacterium pseudotuberculosis* and its role in ovine caseous lymphadenitis. **Journal of comparative pathology**, v. 137, p. 179-210, 2007.

BARAKAT AA, SELIM AS, ATEF A, SABER MS, NAFIE EK, EL-EBEEDY AA. Two serotypes of *Corynebacterium pseudotuberculosis* isolated from different animal species. **Rev. sci. tech. Off. int. Epiz.**, 1984, 3 (1), 151-163

BATEY RG. Pathogenesis of caseous lymphadenitis in sheep and goats. **Aust. Vet. J.** (1986) 63: pp. 269-72.

BAROCCHI, M. A.; RIES, J.; ZOGAJ, X.; HEMSLEY, C.; ALBIGER, B.; KANTH, A.; DAHLBERG, S.; FERNEBRO, J.; MOSCHIONI, M.; MASIGNANI, V. et al. A pneumococcal pilus influences virulence and host inflammatory responses. **Proceedings of the National Academy of Sciences of the United States of America**, v. 103, n. 8, p. 2857-2862, February 21, 2006 2006.

BARRIUSO-IGLESIAS, M.; BARREIRO, C.; FLECHOSO, F.; MARTÍN, J. F. Transcriptional analysis of the FOF1 ATPase operon of *Corynebacterium glutamicum* ATCC 13032 reveals strong induction by alkaline pH. **Microbiology**, v. 152, n. 1, p. 11-21, January 1, 2006 2006.

BEJERANO-SAGIE, M.; XAVIER, K. B. The role of small RNAs in quorum sensing. **Current opinion in microbiology**, v. 10, n. 2, p. 189-198, 2007.

BESSMAN, M. J.; FRICK, D. N.; O'HANDLEY, S. F. The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. **The Journal of biological chemistry**, v. 271, p. 25059-62, 1996.

BESSMAN, M. J.; WALSH, J. D.; DUNN, C. A.; SWAMINATHAN, J.; WELDON, J. E.; SHEN, J. The gene ygdP, associated with the invasiveness of *Escherichia coli* K1, designates a Nudix hydrolase, Orf176, active on adenosine (5')-pentaphospho-(5')-adenosine (Ap5A). **The Journal of biological chemistry**, v. 276, p. 37834-8, 2001.

- BETTS, J. C.; LUKEY, P. T.; ROBB, L. C.; MCADAM, R. A.; DUNCAN, K. Evaluation of a nutrient starvation model of Mycobacterium tuberculosis persistence by gene and protein expression profiling. **Molecular microbiology**, v. 43, n. 3, p. 717-731, 2002.
- BILLINGTON, S. J.; ESMAY, P. A.; SONGER, J. G.; JOST, B. H. Identification and role in virulence of putative iron acquisition genes from Corynebacterium pseudotuberculosis. **FEMS Microbiology Letters**, v. 208, p. 41-45, 2002.
- BINNS S. H., GREEN L. E., BAILEY M. Development and validation of an ELISA to detect antibodies to *Corynebacterium pseudotuberculosis* in ovine sera. *Vet. Microbiol.*, v. 123, p. 169-179, 2007.
- BOCHNER, B. R.; LEE, P. C.; WILSON, S. W.; CUTLER, C. W.; AMES, B. N. AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress. **Cell**, v. 37, n. 1, p. 225-232, 1984.
- BOGDAN, J.; NEWLANDS-MONTEITH, C.; ELLIS, J. Nitric oxide production following in vitro stimulation of ovine pulmonary alveolar macrophages. **Vet Immunol Immunopathol.**, v. 56, p. 299-310, 1997.
- BOREZEE, E.; PELLEGRINI, E.; BERCHE, P. Protein Required for Bacterial Growth at Low Temperature and Involved in Intracellular Survival. **Society**, v. 68, p. 7069-7077, 2000.
- BRANTL, S. Regulatory mechanisms employed by cis-encoded antisense RNAs. **Current opinion in microbiology**, v. 10, n. 2, p. 102-109, 2007.
- BRAITHWAITE CE, SMITH EE, SONGER JG & REINE AH. Characterization of detergent soluble proteins of *corynebacterium pseudotuberculosis*. **Vet. Microbiol** 38: 1993
- BREHMER, D.; GÄSSLER, C.; RIST, W.; MAYER, M. P.; BUKAU, B. Influence of GrpE on DnaK-Substrate Interactions. **Journal of Biological Chemistry**, v. 279, n. 27, p. 27957-27964, July 2, 2004 2004.
- BROGDEN KA, CHEDID L, CUTLIP RC, LEHMKUHL HD, SACKS J. Effect of muramyl dipeptide on immunogenicity of *Corynebacterium pseudotuberculosis* whole-cell vaccines in mice and lambs. **Am. J. Vet. Res.** 51: pp. 200-2. 1990
- BROUNS, S. J. J.; JORE, M. M.; LUNDGREN, M.; WESTRA, E. R.; SLIJKHUIS, R. J. H.; SNIJDERS, A. P. L.; DICKMAN, M. J.; MAKAROVA, K. S.; KOONIN, E. V.; VAN DER OOST, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. **Science (New York, N.Y.)**, v. 321, p. 960-4, 2008.
- BROWN, J. L.; ROSS, T.; MCMEEKIN, T. A.; NICHOLS, P. D. Acid habituation of Escherichia coli and the potential role of cyclopropane fatty acids in low pH tolerance. **International journal of food microbiology**, v. 37, p. 163-73, 1997.
- BUCCA, G.; BRASSINGTON, A. M. E.; HOTCHKISS, G.; MERSINIAS, V.; SMITH, C. P. Negative feedback regulation of dnaK, clpB and lon expression by the DnaK chaperone machine in Streptomyces coelicolor, identified by transcriptome and in vivo DnaK-depletion analysis. **Molecular microbiology**, v. 50, n. 1, p. 153-166, 2003.

BUCHMEIER, N. A.; LIPPS, C. J.; SO, M. Y.; HEFFRON, F. Recombination-deficient mutants of *Salmonella typhimurium* are avirulent and sensitive to the oxidative burst of macrophages. **Mol Microbiol**, v. 7, n. 6, p. 933-6, Mar 1993.

BUCK, G.A., CROSS, R.E., WONG, T.P., LOERA, J., GROMAN, N. DNA relationships among some *tox*-bearing corynebacteriophages. **Infect. Immun.** V.49, p. 679-684, 1985.

BUKAU, B. Regulation of the *Escherichia coli* heat-shock response. **Molecular microbiology**, v. 9, n. 4, p. 671-680, 1993.

BUNIK, V.; SCHLOSS, J.; PINTO, J.; GIBSON, G.; COOPER, A. Enzyme-Catalyzed Side Reactions with Molecular Oxygen may Contribute to Cell Signaling and Neurodegenerative Diseases. **Neurochemical Research**, v. 32, n. 4, p. 871-891, 2007.

CALAMITA H, KO C, TYAGI S, YOSHIMATSU T, MORRISON NE, BISHAI WR. The *Mycobacterium tuberculosis* SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence. **Cell Microbiol.** 7(2):233-44. 2005

CAMARENA, L.; BRUNO, V.; EUSKIRCHEN, G.; POGGIO, S.; SNYDER, M. Molecular mechanisms of ethanol-induced pathogenesis revealed by RNA-sequencing. **PLoS pathogens**, v. 6, p. e1000834, 2010.

CASES I, DE LORENZO V. Promoters in the environment: transcriptional regulation in its natural context. **Nat Rev Microbiol.** 3(2):105-18. 2005

CASTRO TLP. **Avaliação da expressão diferencial de genes codificadores de fatores sigma de *Corynebacterium pseudotuberculosis* em resposta a agentes geradores de estresse oxidativo.** Dissertação (Mestrado em Genética) -Universidade Federal de Minas Gerais, BH, MG (2009): pp. 1-112.

CETINKAYA, B., KARAHAN, M., ATIL, E., KALIN, R., DE BAERE, T., VENEECHOUTTE, M. Identification of *Corynebacterium pseudotuberculosis* isolates from sheep and goats by PCR. **Vet Microbiol.** 88: 75-83. 2002

CHAPLIN PJ, DE ROSE R, BOYLE JS, MCWATERS P, KELLY J, TENNENT JM, LEW AM, SCHEERLINCK JP. Targeting improves the efficacy of a Dna vaccine against *Corynebacterium pseudotuberculosis* in sheep. **Infect. Immun.** 67: pp. 6434-8. 1999

COELHO, K.S. **Isolamento, clonagem e caracterização molecular do gene hsp60 de *Corynebacterium pseudotuberculosis* e sua utilização na construção de uma vacina de subunidade protéica e de DNA.** Dissertação (Mestrado em Genética)-Universidade Federal de Minas Gerais, BH, MG (2007): p. 1-165

COKER C, BAKARE OO, MOBLEY HL. H-NS is a repressor of the *Proteus mirabilis* urease transcriptional activator gene ureR. **J Bacteriol** 2000 182(9):2649-53.

COLLETT MG, BATH GF, CAMERON CM. *Corynebacterium pseudotuberculosis* infections. in: infections diseases of livestock with special reference to southern africa. **Oxford University Press** 2: pp. 1387-95. 1994

CORNELIS, G.; SLUITERS, C.; DE ROUVROIT, C. L.; MICHIELS, T. Homology between virF, the transcriptional activator of the Yersinia virulence regulon, and AraC, the Escherichia coli arabinose operon regulator. **J. Bacteriol.**, v. 171, n. 1, p. 254-262, January 1, 1989 1989.

COSSART, P.; JONQUIÈRES, R. Sortase, a universal target for therapeutic agents against gram-positive bacteria? **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, p. 5013-5, 2000.

COSTERTON JW, STEWART PS, GREENBERG EP. Bacterial biofilms: a common cause of persistent infections. **Science** 284: pp. 318–322.1999

COWDHURY R, SAHU GMK & DAS J. Stress response in pathogenic bacteria. **J. Biosci.** 21: p. 149–160. 1996

CRAIG; GAMBILL; NELSON, A. Heat shock proteins: molecular chaperones of protein biogenesis. **Microbiol Rev**, v. 57(2), p. 402–414, 1993.

CRAIG JE, NOBBS A, HIGH NJ The extracytoplasmic sigma factor, final sigma(E), is required for intracellular survival of nontypeable *Haemophilus influenzae* in J774 macrophages. **Infect Immun.** (2002) 70(2):708-15.

CRONAN, J. E., JR. Phospholipid Alterations During Growth of Escherichia coli. **J. Bacteriol.**, v. 95, n. 6, p. 2054-2061, June 1, 1968 1968.

CSONKA LN. Physiological and genetic responses of bacteria to osmotic stress. **Microbiol Rev**, v.53, p. 121-47, 1989

CUI, S.; MENG, J.; BHAGWAT, A. A. Availability of Glutamate and Arginine during Acid Challenge Determines Cell Density-Dependent Survival Phenotype of Escherichia coli Strains. **Appl. Environ. Microbiol.**, v. 67, n. 10, p. 4914-4918, October 1, 2001 2001.

CUMMINGS, H. S.; HERSHEY, J. W. Translation initiation factor IF1 is essential for cell viability in Escherichia coli. **Journal of bacteriology**, v. 176, p. 198-205, 1994.

CUSSAC V, FERRERO RL, LABIGNE A. Expression of *Helicobacter pylori* urease genes in Escherichia coli grown under nitrogen-limiting conditions. **J. Bacteriol** (1992)174(8):2466-73

DAHLQUIST, K. D.; PUGLISI, J. D. Interaction of translation initiation factor IF1 with the E. coli ribosomal A site. **Journal of molecular biology**, v. 299, p. 1-15, 2000.

DALEBROUX, Z. D.; YAGI, B. F.; SAHR, T.; BUCHRIESER, C.; SWANSON, M. S. Distinct roles of ppGpp and DksA in Legionella pneumophila differentiation. **Molecular microbiology**, v. 76, n. 1, p. 200-219, 2010.

DAS GUPTA, T.; BANDYOPADHYAY, B.; DAS GUPTA, S. K. Modulation of DNA-binding activity of Mycobacterium tuberculosis HspR by chaperones. **Microbiology (Reading, England)**, v. 154, p. 484-90, 2008.

DAS, S.; KANAMOTO, T.; GE, X.; XU, P.; UNOKI, T.; MUNRO, C. L.; KITTEN, T. Contribution of Lipoproteins and Lipoprotein Processing to Endocarditis Virulence in *Streptococcus sanguinis*. **J. Bacteriol.**, v. 191, n. 13, p. 4166-4179, July 1, 2009 2009.

DEDECEK, A.S. **Análise da ilha genômica hrp/hrc de *Herbaspirillum seropedicae* e caracterização parcial dos genes hrcC, hrcV e hrpG.**, Dissertação (Mestrado em Ciências-Bioquímica) - Universidade Federal do Paraná, Curitiba, 2006: p. 85

DELAYE, L.; BECERRA, A.; ORGEL, L.; LAZCANO, A. Molecular Evolution of Peptide Methionine Sulfoxide Reductases (MsrA and MsrB): On the Early Development of a Mechanism That Protects Against Oxidative Damage. **Journal of Molecular Evolution**, v. 64, n. 1, p. 15-32, 2007.

DERCKSEN, D.P.; BRINKHOF, J.M.A.; DEKKER- NOOREN, T.; VAN MAANEN, K.; BODE, C.F.; BAIRD, G.; KAMP, E.M. A comparison of four serological tests for the diagnosis of caseous lymphadenitis in sheep and goats. **Vet. Microbiol**, v. 75, p. 167–175, 2000.

DHANDAYUTHAPANI, S.; JAGANNATH, C.; NINO, C.; SAIKOLAPPAN, S.; SASINDRAN, S. J. Methionine sulfoxide reductase B (MsrB) of *Mycobacterium smegmatis* plays a limited role in resisting oxidative stress. **Tuberculosis (Edinburgh, Scotland)**, v. 89 Suppl 1, p. S26-32, 2009.

DIAS, A. A. S. O.; SILVA, F. C.; PEREIRA, G. A.; SOUZA, M. C.; CAMELLO, T. C. F.; DAMASCENO, J. A. L. D.; PACHECO, L. G. C.; MIYOSHI, A.; AZEVEDO, V. A.; HIRATA, R. et al. *Corynebacterium ulcerans* isolated from an asymptomatic dog kept in an animal shelter in the metropolitan area of Rio de Janeiro, Brazil. **Vector borne and zoonotic diseases (Larchmont, N.Y.)**, v. 10, p. 743-8, 2010.

DIMROTH P, KAIM G, MATTHEY U. Crucial role of the membrane potential for ATP synthesis by F(1)F(o) ATP synthases. **J Exp Biol**. 203(Pt 1):51-9. 2000.

DOMINGUETTI CP. **Análise do papel do fator *sigma C* na resposta de *Corynebacterium pseudotuberculosis* a diferentes condições de estresse ambiental.** - Dissertação (Mestrado em Genética)- Universidade Federal de Minas Gerais, BH, MG, (2011):pp 1-114.

DORELLA, F.A. Análise do potencial vacinal de linhagens recombinantes e selvagens inativadas de *Corynebacterium pseudotuberculosis*. Tese (Doutorado em Genética)- Universidade Federal de Minas Gerais, BH, MG (2009): pp.1-191.

DORELLA, F. A.; PACHECO, L. G.; OLIVEIRA, S. C.; MIYOSHI, A.; AZEVEDO, V. *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. **Vet Res**, v. 37, n. 2, p. 201-18, 2006a

DORELLA FA, ESTEVAM EM, PACHECO LGC, GUIMARÃES CT, LANA UGP, GOMES EA, BARSANTE MM, OLIVEIRA SC, MEYER R, MIYOSHI A e AZEVEDO V. In vivo insertional mutagenesis in *Corynebacterium pseudotuberculosis*: an efficient means to identify dna sequences encoding exported proteins. **Appl. Environ. Microbiol**. 72: pp. 7368-72. 2006b

- DORELLA FA, PACHECO LG, SEYFFERT N, PORTELA RW, MEYER R, MIYOSHI A e AZEVEDO V. Antigens of *Corynebacterium pseudotuberculosis* and prospects for vaccine development. **Expert Rev Vaccines** 8: pp. 205-13. 2009
- DORMAN, C. J. DNA supercoiling and environmental regulation of gene expression in pathogenic bacteria. **Infect. Immun.**, v. 59, n. 3, p. 745-749, March 1, 1991 1991.
- DOUKHAN L, PREDICH M, NAIR G, DUSSURGET O, MANDIC-MULEC I, COLE ST, SMITH DR, SMITH I. Genomic organization of the mycobacterial sigma gene cluster. **Gene**. (1995)165(1):67-70.
- DOW JM & DANIELS MJ. Xylella genomics and bacterial pathogenicity to plants. **Yeast** 17: pp. 263-71. 2000
- DRAZEK ES, HAMMACK CA, SCHMITT MP. *Corynebacterium diphtheriae* genes required for acquisition of iron from haemin and haemoglobin are homologous to ABC haemin transporters. **Mol Microbiol**. 2000 36(1):68-84.
- D'SOUZA, S.; ROSSEELS, V.; DENIS, O.; TANGHE, A.; DE SMET, N.; JURION, F.; PALFLIET, K.; CASTIGLIONI, N.; VANONCKELEN, A.; WHEELER, C. et al. Improved Tuberculosis DNA Vaccines by Formulation in Cationic Lipids. **Infect. Immun.**, v. 70, n. 7, p. 3681-3688, 2002.
- D'SOUZA-AULT MR, SMITH LT, SMITH GM. Roles of N-acetylglutaminylglutamine amide and glycine betaine in adaptation of *Pseudomonas aeruginosa* to osmotic stress. **Appl Environ Microbiol**.v. 59, p.473-8, 1993.
- DUBIN, G. Extracellular Proteases of *Staphylococcus* spp. **Biological Chemistry**, v. 383, n. 7-8, p. 1075-1086, 2002/08/01 2002.
- DUBNAU, E.; LANÉELLE, M. A.; SOARES, S.; BÉNICHOU, A.; VAZ, T.; PROMÉ, D.; PROMÉ, J. C.; DAFFÉ, M.; QUÉMARD, A. Mycobacterium bovis BCG genes involved in the biosynthesis of cyclopropyl keto- and hydroxy-mycolic acids. **Molecular microbiology**, v. 23, p. 313-22, 1997.
- DUNDERDALE, H. J.; BENSON, F. E.; PARSONS, C. A.; SHARPIES, G. J.; LLOYD, R. G.; WEST, S. C. Formation and resolution of recombination intermediates by E. coli RecA and RuvC proteins. **Nature**, v. 354, n. 6354, p. 506-510, 1991.
- DUROCHER, D.; TAYLOR, I. A.; SARBASSOVA, D.; HAIRE, L. F.; WESTCOTT, S. L.; JACKSON, S. P.; SMERDON, S. J.; YAFFE, M. B. The Molecular Basis of FHA Domain: Phosphopeptide Binding Specificity and Implications for Phospho-Dependent Signaling Mechanisms. **Molecular Cell**, v. 6, n. 5, p. 1169-1182, 2000.
- EDELSTEIN, P. H.; HU, B.; SHINZATO, T.; EDELSTEIN, M. A. C.; XU, W.; BESSMAN, M. J. Legionella pneumophila NudA Is a Nudix Hydrolase and Virulence Factor. **Society**, v. 73, p. 6567-6576, 2005.

EGEN NB, CUEVAS WA, MCNAMARA PJ, SAMMONS DW, HUMPHREYS R, SONGER JG. Purification of the phospholipase D of *Corynebacterium pseudotuberculosis* by recycling isoelectric focusing. **Am. J. Vet. Res.** 50: pp. 1319-22. 1989

EGGLETON DG, MIDDLETON HD, DOIDGE CV, MINTY DW. Immunisation against ovine caseous lymphadenitis: comparison of *Corynebacterium pseudotuberculosis* vaccines with and without bacterial cells. **Aust. Vet. J.** 68: pp. 317-9. 1991

EHIRA S, TERAMOTO H, INUI M, YUKAWA H. Regulation of *Corynebacterium glutamicum* heat shock response by the extracytoplasmic-function sigma factor SigH and transcriptional regulators HspR and HrcA. **J Bacteriol.** (9):2964-72. 2009

ELSNER, A.; KREIKEMEYER, B.; BRAUN-KIEWNICK, A.; SPELLERBERG, B.; BUTTARO, B. A.; PODBIELSKI, A. Involvement of Lsp, a Member of the LraI-Lipoprotein Family in *Streptococcus pyogenes*, in Eukaryotic Cell Adhesion and Internalization. **Infect. Immun.**, v. 70, n. 9, p. 4859-4869, 2002

ENGELS S, SCHWEITZER JE, LUDWIG C, BOTT M, SCHAFFER S. clpC and clpP1P2 gene expression in *Corynebacterium glutamicum* is controlled by a regulatory network involving the transcriptional regulators ClgR and HspR as well as the ECF sigma factor sigma H. **Mol Microbiol.** 52(1):285-302.2004.

FARR, S.; KOGOMA, T. Oxidative stress responses in *Escherichia coli* and *Salmonella typhimurium*. **Microbiol Rev.**, v. 55(4), p. 561-585, 1991.

FENIOUK BA, JUNGE W. Regulation of the F0F1-ATP synthase: the conformation of subunit epsilon might be determined by directionality of subunit gamma rotation. **FEBS Lett.** (2005) 579(23):5114-8

FENTON, W.; HORWICH, A. Chaperonin-mediated protein folding: fate of substrate polypeptide. **Q Rev Biophys**, v. 36, p. 229-56, 2003.

FERNÁNDEZ EP, VELA AI, LAS HERAS A, DOMÍNGUEZ L, FERNÁNDEZ-GARAYZÁBAL JF e MORENO MA. Antimicrobial susceptibility of *Corynebacteria* isolated from ewe's mastitis. **Int. J. Antimicrob. Agents** (2001) 18: pp. 571-4.

FERREIRA A, GRAY M, WIEDMANN M, BOOR KJ. Comparative genomic analysis of the sigB operon in *Listeria monocytogenes* and in other Gram-positive bacteria. **Curr Microbiol.** 2004 48(1):39-46

FERREIRA A, O'BYRNE CP, BOOR KJ. Role of sigma(B) in heat, ethanol, acid, and oxidative stress resistance and during carbon starvation in *Listeria monocytogenes*. **Appl Environ Microbiol.** (2001) 67(10):4454-7.

FERREIRA A, SUE D, O'BYRNE CP, BOOR KJ. Role of *Listeria monocytogenes* sigma(B) in survival of lethal acidic conditions and in the acquired acid tolerance response. **Appl Environ Microbiol.** (2003) 69(5):2692-8.

FERRI-FIONI, M. L.; SCHMITT, E.; SOUTOURINA, J.; PLATEAU, P.; MECHULAM, Y.; BLANQUET, S. Structure of crystalline D-Tyr-tRNA(Tyr) deacylase. A representative of a

new class of tRNA-dependent hydrolases. **The Journal of biological chemistry**, v. 276, p. 47285-90, 2001.

FINLAY, B. B.; FALKOW, S. Common themes in microbial pathogenicity revisited. **Microbiology and molecular biology reviews : MMBR**, v. 61, p. 136-69, 1997.

FOL, M.; CHAUHAN, A.; NAIR, N. K.; MALONEY, E.; MOOMEY, M.; JAGANNATH, C.; MADIRAJU, M. V. V. S.; RAJAGOPALAN, M. Modulation of Mycobacterium tuberculosis proliferation by MtrA, an essential two-component response regulator. **Molecular microbiology**, v. 60, n. 3, p. 643-657, 2006.

FOZO, E. M.; QUIVEY, R. G., JR. The fabM Gene Product of Streptococcus mutans Is Responsible for the Synthesis of Monounsaturated Fatty Acids and Is Necessary for Survival at Low pH. **J. Bacteriol.**, v. 186, n. 13, p. 4152-4158, July 1, 2004 2004.

FRANCIS JR RT, BOOTH JW, BECKER RR. Uptake of iron from hemoglobin and the haptoglobin-hemoglobin complex by hemolytic bacteria. **Int J Biochem** 17: 767-73. 1985

FUTAI, M., NOUMI, T., AND MAEDA, M. ATP synthase (H⁺-ATPase): results by combined biochemical and molecular biological approaches. **Annu. Rev. Biochem.** 58, 111-13. 1989

FROTA, C. C.; PAPA VINASASUNDARAM, K. G.; DAVIS, E. O.; COLSTON, M. J. The AraC Family Transcriptional Regulator Rv1931c Plays a Role in the Virulence of Mycobacterium tuberculosis. **Infect. Immun.**, v. 72, n. 9, p. 5483-5486. 2004.

FUGE, E. K.; FARR, S. B. AppppA-binding protein E89 is the Escherichia coli heat shock protein ClpB. **J. Bacteriol.**, v. 175, n. 8, p. 2321-2326, 1993.

FUTAI, M. N., T.; AND MAEDA, M. ATP synthase (H⁺-ATPase): results by combined biochemical and molecular biological approaches. **Annu. Rev. Biochem**, v. 58, p. 111-13, 1989.

GAN, K.; GUPTA, S. D.; SANKARAN, K.; SCHMID, M. B.; WU, H. C. Isolation and characterization of a temperature-sensitive mutant of Salmonella typhimurium defective in prolipoprotein modification. **Journal of Biological Chemistry**, v. 268, n. 22, p. 16544-16550, 1993.

GAO, C.-H.; YANG, M.; HE, Z.-G. An ArsR-like transcriptional factor recognizes a conserved sequence motif and positively regulates the expression of phoP in mycobacteria. **Biochemical and Biophysical Research Communications**, v. 411, n. 4, p. 726-731, 2011.

GARAY-ARROYO, A.; COLMENERO-FLORES, J. M.; GARCIARRUBIO, A.; COVARRUBIAS, A. A. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. **J. Biol. Chem**, v. 275, p. 5668-5674, 2000.

GARNER MR, NJAA BL, WIEDMANN M, BOOR KJ. Sigma B contributes to *Listeria monocytogenes* gastrointestinal infection but not to systemic spread in the guinea pig infection model. **Infect Immun.** 74(2):876-86. 2006

GOMEZ M, DOUKHAN L, NAIR G, SMITH I. sigA is an essential gene in *Mycobacterium smegmatis*. **Mol Microbiol.** 29(2):617-28. 1998

GOODELL, E. W.; HIGGINS, C. F. Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli*. **J. Bacteriol.**, v. 169, n. 8, p. 3861-3865, 1987.

GOPHNA; RON. Virulence and the heat shock response. **Int J Med Microbiol.** , v. 292 p. 453-61, 2003.

GOVAN, J.R., DERETIC, V. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. **Microbiol Rev.**, v 60, p.539-74, 1996.

GRAHAM JE, CLARK-CURTISS JE. Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). **Proc Natl Acad Sci U S A.** 96(20):11554-9. 1999

GROISMAN, E. A.; CHIAO, E.; LIPPS, C. J.; HEFFRON, F. *Salmonella typhimurium* phoP virulence gene is a transcriptional regulator. **Proceedings of the National Academy of Sciences**, v. 86, n. 18, p. 7077-7081, 1989.

GROSS. Function and regulation of the heat shock response. In: (Ed.). **Escherichia coli and Salmonella: Cellular and Molecular Biology**,. p.1382–1399. 1996

GÜELL, M.; VAN NOORT, V.; YUS, E.; CHEN, W.-H.; LEIGH-BELL, J.; MICHALODIMITRAKIS, K.; YAMADA, T.; ARUMUGAM, M.; DOERKS, T.; KÜHNER, S. et al. Transcriptome complexity in a genome-reduced bacterium. **Science (New York, N.Y.)**, v. 326, p. 1268-71, 2009.

GUIMARÃES AS, SEYFFERT N, PORTELA RWD, MEYER R, CARMO FB, CRUZ J, LAGE AP, HEINEMANN MB, MIYOSHI A, AZEVEDO V, GOUVEIA AMG. Caseous lymphadenitis in sheep flocks of the state of minas gerais, Brazil: prevalence and management surveys. **Small Ruminants Research**, v.87, p:86-91, 2009

GUISBERT, E.; HERMAN, C.; LU, C. Z.; GROSS, C. A. A chaperone network controls the heat shock response in *E. coli*. **Genes e Development**, v. 18, n. 22, p. 2812-2821, November 15, 2004 2004.

HAFT, D. H.; SELENGUT, J.; MONGODIN, E. F.; NELSON, K. E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. **PLoS computational biology**, v. 1, p. e60, 2005.

HAIKARAINEN, T.; PAPAGEORGIOU, A. C. Dps-like proteins: structural and functional insights into a versatile protein family. **Cellular and molecular life sciences : CMLS**, v. 67, p. 341-51, 2010.

HALDENWANG WG. The sigma factors of *Bacillus subtilis*. **Microbiol Rev** 1995, 59:1-30.

HAMILTON, A.; ROBINSON, C.; SUTCLIFFE, I. C.; SLATER, J.; MASKELL, D. J.; DAVIS-POYNTER, N.; SMITH, K.; WALLER, A.; HARRINGTON, D. J. Mutation of the Maturase Lipoprotein Attenuates the Virulence of *Streptococcus equi* to a Greater Extent than

Does Loss of General Lipoprotein Lipidation. **Infect. Immun.**, v. 74, n. 12, p. 6907-6919, 2006.

HAN, Y.; ZHOU, D.; PANG, X.; ZHANG, L.; SONG, Y.; TONG, Z.; BAO, J.; DAI, E.; WANG, J.; GUO, Z. et al. Comparative transcriptome analysis of *Yersinia pestis* in response to hyperosmotic and high-salinity stress. **Research in microbiology**, v. 156, p. 403-15, 2005.

HANNA, M. N.; FERGUSON, R. J.; LI, Y.-H.; CVITKOVITCH, D. G. *uvrA* Is an Acid-Inducible Gene Involved in the Adaptive Response to Low pH in *Streptococcus mutans*. **J. Bacteriol.**, v. 183, n. 20, p. 5964-5973, 2001.

HANNAPPEL, A.; BUNDSCHUH, F.; GREINER, P.; ALLES, M.; WERNER, C.; RICHTER, O.; LUDWIG, B. bacterial model systems for cytochrome c oxidase biogenesis. **Indian journal of chemistry**, v. 50A, p. 374-382, 2011.

HAYDON, D.; GUEST, J. A new family of bacterial regulatory proteins. **FEMS**, v. 63, p. 291 - 295, 1991.

HECKER, M.; SCHUMANN, W.; VO, U. MicroReview Heat-shock and general stress response in *Bacillus subtilis*. **Molecular Microbiology**, v. 19, p. 417-428, 1996.

HECKER M, VOLKER U: General stress response of *Bacillus subtilis* and other bacteria. **Adv Microb Physiol** 2001, 44:35-91.

HEDENFALK, I.; DUGGAN, D.; CHEN, Y.; RADMACHER, M.; BITTNER, M.; SIMON, R.; MELTZER, P.; GUSTERSON, B.; ESTELLER, M.; RAFFELD, M. et al. Gene-Expression Profiles in Hereditary Breast Cancer. **New England Journal of Medicine**, v. 344, n. 8, p. 539-548, 2001.

HELMANN JD. Sigma Factors in Gene Expression. *eLS*. 2005

HELMANN JD, CHAMBERLIN MJ. Structure and function of bacterial sigma factors. **Annu Rev Biochem.** (1988) 57:839-72.

HENGGE-ARONIS, R. Recent insights into the general stress response regulatory network in *Escherichia coli*. **Journal of molecular microbiology and biotechnology**, v. 4, p. 341-6, 2002.

HEURGUE-HAMARD, V.; CHAMP, S.; ENGSTROM, A.; EHRENBERG, M.; BUCKINGHAM, R. H. The *hemK* gene in *Escherichia coli* encodes the N5-glutamine methyltransferase that modifies peptide release factors. **EMBO J**, v. 21, n. 4, p. 769-778, 2002.

HIGGINS, C. F. ABC transporters: physiology, structure and mechanism--an overview. **Res Microbiol**, v. 152, p. 205-10, 2001.

HILPERT, W.; DIMROTH, P. Purification and Characterization of a New Sodium-Transport Decarboxylase. **European Journal of Biochemistry**, v. 132, n. 3, p. 579-587, 1983.

- HIRAMATSU T, KODAMA K, KURODA T, MIZUSHIMA T, TSUCHIYA T. A putative multisubunit Na⁺/H⁺ antiporter from *Staphylococcus aureus*. **J Bacteriol.** 6642-8. 1998
- HIRSH EC, ZEE YC. **Microbiologia veterinária. Guanabara-Koogan, Rio de Janeiro: Brasil:** p. 446. 2003
- HODGSON AL, CARTER K, TACHEDJIAN M, KRYWULT J, CORNER LA, MCCOLL M, CAMERON A. Efficacy of an ovine caseous lymphadenitis vaccine formulated using a genetically inactive form of the corynebacterium pseudotuberculosis phospholipase d. **Vaccine** 17: pp. 802-8. 1999
- HOLSTAD G. Corynebacterium pseudotuberculosis infection in goats. ix. the effect of vaccination against natural infection. **Acta Vet. Scand.** 30: pp. 285-93. 1989
- HOLT JG, KRIEG NR, SNEATH PHA, STALEY JT e WILLIAM ST. Regular, nonsporing gram-positive rods. bergey's manual of determinative bacteriology, 9th ed. **The Williams e Wilkins Co., Baltimore, Md** pp. 565-7. 1994
- HOPKINS, D. W.; O'DOWD, R. W.; SHIEL, R. S. Comparison of D- and L-amino acid metabolism in soils with differing microbial biomass and activity. **Soil Biology and Biochemistry**, v. 29, p. 23-29, 1997.
- HOSKISSON, P. A.; HUTCHINGS, M. I. MtrAB-LpqB: a conserved three-component system in actinobacteria? **Trends in microbiology**, v. 14, p. 444-9, 2006.
- HUDER, J. B.; DIMROTH, P. Expression of the sodium ion pump methylmalonyl-coenzyme A-decarboxylase from *Veillonella parvula* and of mutated enzyme specimens in *Escherichia coli*. **Journal of bacteriology**, v. 177, p. 3623-30, 1995.
- HUGHES KT, MATHEE K. The anti-sigma factors **Annu Rev Microbiol.** 52:231-86. 1998
- HURST, A. Bacterial injury: a review. **Canadian Journal of Microbiology**, v. 23, p. p.935–944, 1977.
- HUTCHINGS, M. J.; JOHN, E. A.; STEWART, A. J. A. The Ecological Consequences of Environmental Heterogeneity. **Blackwell Science**, 2000.
- IMLAY JA, FRIDOVICH I. Superoxide production by respiring membranes of *Escherichia coli*. **Free Radic Res Commun** 59-66. 1991
- INBAR, O.; RON, E. Induction of cadmium tolerance in *Escherichia coli* K-12. **FEMS Microbiol Lett**, v. 113, p. 197-200, 1993.
- INOUYE, S.; WANG, S.; SEKIZAWA, J.; HALEGOUA, S.; INOUYE, M. Amino acid sequence for the peptide extension on the prolipoprotein of the *Escherichia coli* outer membrane. **Proc Natl Acad Sci USA**, v. 74, p. 1004-8, 1977.
- ISABELLA, V. M.; CLARK, V. L. Deep sequencing-based analysis of the anaerobic stimulon in *Neisseria gonorrhoeae*. **BMC genomics**, v. 12, p. 51, 2011.

ITALIANI VC, DA SILVA NETO JF, BRAZ VS, MARQUES MV. Regulation of catalase-peroxidase KatG is OxyR dependent and Fur independent in *Caulobacter crescentus*. **J Bacteriol.** 193(7):1734-44. 2011

JAKOB, K.; SATORHELYI, P.; LANGE, C.; WENDISCH, V. F.; SILAKOWSKI, B.; SCHERER, S.; NEUHAUS, K. Gene Expression Analysis of *Corynebacterium glutamicum* Subjected to Long-Term Lactic Acid Adaptation. **J. Bacteriol.**, v. 189, n. 15, p. 5582-5590, August 1, 2007 2007.

JANZ, D.; BEHNKE, K.; SCHNITZLER, J.-P.; KANAWATI, B.; SCHMITT-KOPPLIN, P.; POLLE, A. Pathway analysis of the transcriptome and metabolome of salt sensitive and tolerant poplar species reveals evolutionary adaptation of stress tolerance mechanisms. **BMC plant biology**, v. 10, p. 150, 2010.

JEAN D, BRIOLAT V., REYSSET G. Oxidative stress response in *Clostridium perfringens*. **Microbiology** 150: pp. 1649-59. 2004

JENKS, P. J.; KUSTERS, J. G. Pathogenesis and virulence of *Helicobacter pylori*. **curr. Opin. Gastroentero**, 2000.

JOHN, S.; JOHN, S. Strategies to Control Stress- Adapted Pathogens. In: (Ed.). **Microbial Stress Adaptation and Food Safety**: CRC Press, ISBN 978-1-56676-912-9. 2002.

JOIN-LAMBERT OF, OUACHE M, CANIONI D, BERETTI JL, BLANCHE S, BERCHE P, KAYAL S. *Corynebacterium pseudotuberculosis* necrotizing lymphadenitis in a twelve-year-old patient. **Pediatr Infect Dis J.** 25(9):pp. 848-51. 2006

JOLLY RD. Some observations on surface lipids of virulent and attenuated strains of *Corynebacterium* tipo I. **J.Appl. Bacteriol** 29: pp. 189-96. 1966

JOSEFSSON, E.; MCCREA, K. W.; NÍ EIDHIN, D.; O'CONNELL, D.; COX, J.; HÖÖK, M.; FOSTER, T. J. Three new members of the serine-aspartate repeat protein multigene family of *Staphylococcus aureus*. **Microbiology (Reading, England)**, v. 144 (Pt 1, p. 3387-95, 1998.

JOU, R.; COWAN, J. Ribonuclease H activation by inert transition-metal complexes. Mechanistic probes for metallocofactors: Insights on the metallobiochemistry of divalent magnesium ion. **Am Chem Soc**, v. 113, p. 6685–6686., 1991.

JOZEF CZUK, S.; KLIE, S.; CATCHPOLE, G.; SZYMANSKI, J.; CUADROS-INOSTROZA, A.; STEINHAUSER, D.; SELBIG, J.; WILLMITZER, L. Metabolomic and transcriptomic stress response of *Escherichia coli*. **Molecular systems biology**, v. 6, p. 364, 2010.

JUSTICE, S. S.; HUNSTAD, D. A.; HARPER, J. R.; DUGUAY, A. R.; PINKNER, J. S.; BANN, J.; FRIEDEN, C.; SILHAVY, T. J.; HULTGREN, S. J. Periplasmic Peptidyl Prolyl cis-trans Isomerases Are Not Essential for Viability , but SurA Is Required for Pilus Biogenesis in *Escherichia coli*. **Society**, v. 187, p. 7680-7686, 2005.

KALINOWSKI, J., BATHE B, BARTELS D, BISCHOFF N, et al. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. **J Biotechnol.** 104(1-3):5-25. 2003

KANG CM, ABBOTT DW, PARK ST, DASCHER CC, CANTLEY LC, HUSSON RN. The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. **Genes Dev.** 19(14):1692-704. 2005

KARLIN, S.; MRÁZEK, J. Predicted highly expressed genes of diverse prokaryotic genomes. **Journal of bacteriology**, v. 182, p. 5238-50, 2000.

KATO-YAMADA Y, BALD D, KOIKE M, MOTOHASHI K, HISABORI T, YOSHIDA M. Epsilon subunit, an endogenous inhibitor of bacterial F(1)-ATPase, also inhibits F(0)F(1)-ATPase. **J Biol Chem.** 274(48):33991-4.1999

KATO Y, MATSUI T, TANAKA N, MUNHEYUKI E, HISABORI T, YOSHIDA M. Thermophilic F1-ATPase is activated without dissociation of an endogenous inhibitor, epsilon subunit. **J Biol Chem.** 272(40):24906-12. 1997

KATO-YAMADA, Y.; YOSHIDA, M.; HISABORI, T. Movement of the Helical Domain of the ϵ Subunit Is Required for the Activation of Thermophilic F1-ATPase. **Journal of Biological Chemistry**, v. 275, n. 46, p. 35746-35750, November 17, 2000

KAUSHAL D, SCHROEDER BG, TYAGI S, YOSHIMATSU T, et al. Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH. **Proc Natl Acad Sci USA** 99: 8330–8335. 2002

KAZMIERCZAK MJ, MITHOE SC, BOOR KJ, WIEDMANN M. *Listeria monocytogenes* sigma B regulates stress response and virulence functions. **J. Bacteriol.** 185: pp. 5722-34. 2003.

KAZMIERCZAK MJ, WIEDMANN M, BOOR KJ. Alternative sigma factors and their roles in bacterial virulence. **Microbiol Mol Biol Rev** 69(4):527-43. 2005

KELLER, P.; BOTTGER, E.; SANDER, P. Tuberculosis vaccine strain *Mycobacterium bovis* BCG Russia is a natural recA mutant. **BMC Microbiology**, v. 8, n. 1, p. 120, 2008.

KENDALL, S. L.; WITHERS, M.; SOFFAIR, C. N.; MORELAND, N. J.; GURCHA, S.; SIDDEERS, B.; FRITA, R.; TEN BOKUM, A.; BESRA, G. S.; LOTT, J. S. et al. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. **Molecular microbiology**, v. 65, p. 684-99, 2007.

KHAMIS, A.; RAOULT, D.; LA SCOLA, B. rpoB gene sequencing for identification of *Corynebacterium* species. **J. Clin. Microbiol.**, v. 42, p. 3925–3931. 2004.

KIM, B.J.; LEE, S.H.; LYU, M.A.; KIM, S.J.; BAI, G.H.; KIM, S.J.; CHAE, G.T.; KIM, E.C.; CHA, C.Y.; KOOK, Y.H. Identification of *Mycobacterium* species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). **J. Clin. Microbiol.**, v. 37, p. 1714–1720. 1999.

KIRK JS, SUMNER JB. Antiurease. **J. Biol. Chem.** 94:21–28. 1934

KISSELEV, L. L.; BUCKINGHAM, R. H. Translational termination comes of age. **Trends in Biochemical Sciences**, v. 25, n. 11, p. 561-566, 2000.

KOIDE, T.; REISS, D. J.; BARE, J. C.; PANG, W. L.; FACCIOTTI, M. T.; SCHMID, A. K.; PAN, M.; MARZOLF, B.; VAN, P. T.; LO, F.-Y. et al. Prevalence of transcription promoters within archaeal operons and coding sequences. **Molecular systems biology**, v. 5, p. 285, 2009.

KOVACIKOVA G, SKORUPSKI K. The alternative sigma factor sigma(E) plays an important role in intestinal survival and virulence in *Vibrio cholerae*. **Infect Immun.** 70(10):5355-62. 2002

KOVACS-SIMON, A.; TITBALL, R. W.; MICHELL, S. L. Lipoproteins of Bacterial Pathogens. **Infect. Immun.**, v. 79, n. 2, p. 548-561, February 1, 2011 2011.

LADEIRA, M. S. P.; SALVADORI, D. M. F.; RODRIGUES, M. A. M. Biopatologia do *Helicobacter pylori*. **Jornal Brasileiro de Patologia Medica e Laboratorial**, v. 39, p. 335-342,, 2003.

LANDAIS, S.; GOUNON, P.; LAURENT-WINTER, C.; MAZIÉ, J. C.; DANCHIN, A.; BÂRZU, O.; SAKAMOTO, H. Immunochemical analysis of UMP kinase from *Escherichia coli*. **Journal of bacteriology**, v. 181, p. 833-40, 1999.

LARISCH C, NAKUNST D, HÜSER AT, TAUCH A, KALINOWSKI J. The alternative sigma factor SigB of *Corynebacterium glutamicum* modulates global gene expression during transition from exponential growth to stationary phase. **BMC Genomics.** 8:4. 2007

LE GUEN, L.; SANTOS, R.; CAMADRO, J. Functional analysis of the hemK gene product involvement in protoporphyrinogen oxidase activity in yeast. **FEMS Microbiol Lett**, v. 173, p. 175-82, 1999.

LEAMASTER, BR., SHEN, DT., GORHAM, JR., LEATHERS, CW., WELLS, HD. Efficacy of *Corynebacterium pseudotuberculosis* bacterin for the immunologic protection of sheep against development of caseous lymphadenitis. **Am. J. Vet. Res.** 48: pp. 869-72. 1987

LEE, E.; AHN, S.; PARK, J.; LEE, J.; AHN, S.; KONG, I. Identification of oligopeptide permease (opp) gene cluster in *Vibrio fluvialis* and characterization of biofilm production by oppA knockout mutation. **FEMS Microbiol Lett**, v. 240, p. 21-30, 2004.

LEIMEISTER-WÄCHTER, M.; DOMANN, E.; CHAKRABORTY, T. The expression of virulence genes in *Listeria monocytogenes* is thermoregulated. **Journal of bacteriology**, v. 174, p. 947-52, 1992.

LIBEREK, K.; WALL, D.; GEORGOPOULOS, C. The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. **Proceedings of the National Academy of Sciences**, v. 92, n. 14, p. 6224-6228, July 3, 1995 1995.

LITWIN CM, CALDERWOOD SB. Role of iron in regulation of virulence genes. **Clin Microbiol Rev.** 6(2):137-49. 1993

- LIU, W.-T.; KARAVOLOS, M. H.; BULMER, D. M.; ALLAOUI, A.; HORMAECHE, R. D. C. E.; LEE, J. J.; KHAN, C. M. A. Role of the universal stress protein UspA of Salmonella in growth arrest, stress and virulence. **Microbial pathogenesis**, v. 42, p. 2-10, 2007.
- LLOYD, R. G. Conjugational recombination in resolvase-deficient *ruvC* mutants of Escherichia coli K-12 depends on *recG*. **J Bacteriol**, v. 173, n. 17, p. 5414-8, Sep 1991.
- LONETTO M, GRIBSKOV M, GROSS CA The sigma 70 family: sequence conservation and evolutionary relationships **J Bacteriol**. 174(12):3843-9. 1992
- LOS, D. A.; MURATA, N. Membrane fluidity and its roles in the perception of environmental signals. **Biochimica et biophysica acta**, v. 1666, p. 142-57, 2004.
- LOUGHLIN, M. F.; BARNARD, F. M.; JENKINS, D.; SHARPLES, G. J.; JENKS, P. J. Helicobacter pylori Mutants Defective in RuvC Holliday Junction Resolvase Display Reduced Macrophage Survival and Spontaneous Clearance from the Murine Gastric Mucosa. **Society**, v. 71, p. 2022-2031, 2003.
- MADDOCKS, S. E.; OYSTON, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. **Microbiology (Reading, England)**, v. 154, p. 3609-23, 2008.
- MAHAN, M.; SLAUCH, J.; MEKALANOS, J. Selection of bacterial virulence genes that are specifically induced in host tissues. **Science**, v. 259, p. 686-8, 1993.
- MAIDEN, M. C. J. Horizontal Genetic Exchange, Evolution, and Spread of Antibiotic Resistance in Bacteria. **Clinical Infectious Diseases**, v. 27, n. Supplement 1, p. S12-S20, 1998.
- MANE, S. P.; EVANS, C.; COOPER, K. L.; CRASTA, O. R.; FOLKERTS, O.; HUTCHISON, S. K.; HARKINS, T. T.; THIERRY-MIEG, D.; THIERRY-MIEG, J.; JENSEN, R. V. Transcriptome sequencing of the Microarray Quality Control (MAQC) RNA reference samples using next generation sequencing. **BMC genomics**, v. 10, p. 264, 2009.
- MANGANELLI R, DUBNAU E, TYAGI S, KRAMER FR, SMITH I. Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. **Mol Microbiol** 31:715–724. 1999
- MANGANELLI R, VOSKUIL MI, SCHOOLNIK GK, SMITH I. The *Mycobacterium tuberculosis* ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. **Mol Microbiol**. 41(2):423-37. 2001
- MARDIS, E. R. The impact of next-generation sequencing technology on genetics. **Trends in genetics : TIG**, v. 24, p. 133-41, 2008a.
- MARDIS, E. R. Next-generation DNA sequencing methods. **Annual review of genomics and human genetics**, v. 9, p. 387-402, 2008b.
- MARLES-WRIGHT J, LEWIS RJ. Stress responses of bacteria. **Curr Opin Struct Biol**. 17(6):755-60. 2007

MARTÍN-GALIANO, A. J.; FERRÁNDIZ, M. J.; DE LA CAMPA, A. G. The promoter of the operon encoding the FOF1 ATPase of *Streptococcus pneumoniae* is inducible by pH. **Molecular microbiology**, v. 41, p. 1327-38, 2001.

MARTIN, J.; ZHU, W.; PASSALACQUA, K. D.; BERGMAN, N.; BORODOVSKY, M. *Bacillus anthracis* genome organization in light of whole transcriptome sequencing. **BMC bioinformatics**, v. 11 Suppl 3, p. S10, 2010.

MATRIX2png. Disponível em <<http://www.bioinformatics.ubc.ca/matrix2png>>. Acessado em 10 de Março, 2011.

MCDEVITT, D.; FRANCOIS, P.; VAUDAUX, P.; FOSTER, T. J. Molecular characterization of the clumping factor (fibrinogen receptor) of *Staphylococcus aureus*. **Molecular microbiology**, v. 11, n. 2, p. 237-248, 1994.

MCGARRITY JT, A. J. The effect of temperature and other growth conditions on the fatty acid composition of *Escherichia coli*. **Can J. Microbiol**, v. 27, p. 835-40, 1981.

MCGEE DJ, MAY CA, GARNER RM, HIMPSL JM, MOBLEY HL. Isolation of *Helicobacter pylori* genes that modulate urease activity. **J Bacteriol**. 8:2477-84. (1999)

MCKEAN, S.; DAVIES, J.; MOORE, R. Expression of phospholipase D the major virulence factor of *Corynebacterium pseudotuberculosis*, is regulated by multiple environmental factors and plays a role in macrophage death. **Microbiology**, v. 153, p. 2203 - 2211, 2007a.

MCKEAN, S. C.; DAVIES, J. K.; MOORE, R. J. Probing the heat shock response of *Corynebacterium pseudotuberculosis* : The major virulence factor , phospholipase D , is downregulated at 43 C. **Research in Microbiology**, v. 158, 2007b.

MCNAMARA, P.; BRADLEY, G.; SONGER, J. Targeted mutagenesis of the phospholipase D gene results in decreased virulence of *Corynebacterium pseudotuberculosis*. **Mol Microbiol**, v. 12, p. 921 - 930, 1994.

MCNAMARA, P. J.; CUEVAS, W. A.; SONGER, J. G. Toxic phospholipases D of *Corynebacterium pseudotuberculosis*, *C. ulcerans* and *Arcanobacterium haemolyticum*: cloning and sequence homology. **Gene**, v. 156, n. 1, p. 113-118, 1995.

MENICHE, X.; LABARRE, C.; DE SOUSA-D'AURIA, C.; HUC, E.; LAVAL, F.; TROPIS, M.; BAYAN, N.; PORTEVIN, D.; GUILHOT, C.; DAFFE, M. et al. Identification of a Stress-Induced Factor of *Corynebacterineae* That Is Involved in the Regulation of the Outer Membrane Lipid Composition. **J. Bacteriol.**, v. 191, n. 23, p. 7323-7332, 2009.

MENZIES PI, HWANG Y.; PRESCOTT JF. Comparison of an interferon-gamma to a phospholipase D enzyme-linked immunosorbent assay for diagnosis of *Corynebacterium pseudotuberculosis* infection in experimentally infected goats. **Vet. Microbiol**. 100: pp. 129-37.2004

MEYER, R.; CARMINATI, R.; BAHIA, R.; VALE, V.; VIEGAS, S.; MARTINEZ, T.; NASCIMENTO, I.; SCHAER, R.; SILVA, J.; RIBEIRO, M. et al. Evaluation of the goats

humoral immune response induced by the *Corynebacterium pseudotuberculosis* lyophilized live vaccine. **J Med Biol Sci**, v. 1, p. 42 - 48, 2002.

MEYER, S. E.; GRANETT, S.; JUNG, J. U.; VILLAREJO, M. R. Osmotic regulation of PhoE porin synthesis in *Escherichia coli*. **J. Bacteriol.**, v. 172, n. 9, p. 5501-5502, 1990

MIETZNER TA, MORSE SA. The role of iron-binding proteins in the survival of pathogenic bacteria. **Annu Rev Nutr.** 14:471-93. 1994

MILLER, J.; MEKALANOS, J.; FALKOW, S. Coordinate regulation and sensory transduction in the control of bacterial virulence. **Science**, v. 243, p. 916-22, 1989.

MOBLEY HL, ISLAND MD, HAUSINGER RP. Molecular biology of microbial ureases. **Microbiol Rev.** 59(3):451-80. 1995

MÖKER, N.; BROCKER, M.; SCHAFFER, S.; KRÄMER, R.; MORBACH, S.; BOTT, M. Deletion of the genes encoding the MtrA–MtrB two-component system of *Corynebacterium glutamicum* has a strong influence on cell morphology, antibiotics susceptibility and expression of genes involved in osmoprotection. **Molecular microbiology**, v. 54, n. 2, p. 420-438, 2004.

MONNET, V. Bacterial oligopeptide-binding proteins. **Cell. Mol. Life Sci**, v. 60 p. 2100–2114, 2003.

MOONEY RA, DARST SA, LANDICK R. Sigma and RNA polymerase: an on-again, off-again relationship? **Mol Cell** 20:335-345. 2005

MOODY, D. E. Genomics techniques : An overview of methods for the study of gene expression The online version of this article , along with updated information and services , is located on the World Wide Web at : Genomics techniques : An overview of methods for the stu. **Journal of Animal Science**, 2001.

MORA, M.; BENSI, G.; CAPO, S.; FALUGI, F.; ZINGARETTI, C.; MANETTI, A. G. O.; MAGGI, T.; TADDEI, A. R.; GRANDI, G.; TELFORD, J. L. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 43, p. 15641-15646, 2005.

MORIMOTO, R. I. Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. **Genes e Development**, v. 12, n. 24, p. 3788-3796, 1998.

MORITA, M.; KANEMORI, M.; YANAGI, H.; YURA, T. Heat-Induced Synthesis of sigma 32 in *Escherichia coli*: Structural and Functional Dissection of rpoH mRNA Secondary Structure. **J. Bacteriol.**, v. 181, n. 2, p. 401-410, 1999.

MOROZOVA, O.; MARRA, M. A. Applications of next-generation sequencing technologies in functional genomics. **Genomics**, v. 92, p. 255-64, 2008.

MORTAZAVI, A.; WILLIAMS, B. A.; MCCUE, K.; SCHAEFFER, L.; WOLD, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, p. 1-8, 2008.

MOSKOVITZ, J.; RAHMAN, M.; STRASSMAN, J.; YANCEY, S.; KUSHNER, S.; BROTH, N.; WEISSBACH, H. Escherichia coli peptide methionine sulfoxide reductase gene: regulation of expression and role in protecting against oxidative damage. **J. Bacteriol.**, v. 177, n. 3, p. 502-507, 1995.

MUCKLE, C. A.; GYLES, C. L. Characterization of strains of corynebacterium pseudotuberculosis. **Canadian journal of comparative medicine. Revue canadienne de médecine comparée**, v. 46, p. 206-8, 1982.

NADON CA, BOWEN BM, WIEDMANN M, BOOR KJ. Sigma B contributes to PrfA-mediated virulence in Listeria monocytogenes. **Infect Immun.** 70(7):3948-52. (2002)

NAGEL, G.; LAHRZ, A.; DERSCH, P. Environmental control of invasins expression in Yersinia pseudotuberculosis is mediated by regulation of RovA, a transcriptional activator of the SlyA/Hor family. **Molecular microbiology**, v. 41, p. 1249-69, 2001.

NAKAYASHIKI, T.; NISHIMURA, K.; INOKUCHI, H. Cloning and sequencing of a previously unidentified gene that is involved in the biosynthesis of heme in Escherichia coli. **Gene**, v. 153, n. 1, p. 67-70, 1995.

NAKUNST D, LARISCH C, HÜSER AT, TAUCH A, PÜHLER A, KALINOWSKI J . The extracytoplasmic function-type sigma factor SigM of *Corynebacterium glutamicum* ATCC 13032 is involved in transcription of disulfide stress-related genes. **J Bacteriol.** 189(13):4696-707. 2007

NARBERHAUS, F.; WALDMINGHAUS, T.; CHOWDHURY, S. RNA thermometers. **FEMS microbiology reviews**, v. 30, n. 1, p. 3-16, 2006.

NCBI. National Center for Biotechnology Information. Disponível em (<http://www.ncbi.nlm.nih.gov/genomes>). Acessado em 28 de Jun, 2011.

NCBI. National Center for Biotechnology Information. Disponível em (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=c109100>). Acessado em 30 Jun, 2011.

NCBI. National Center for Biotechnology Information. Disponível em (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=196798>). Acessado em 30 Jun, 2011.

NCBI. National Center for Biotechnology Information. Disponível em (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=73147>). Acessado em 02 Jul, 2011.

NEPOMUCENO, R. S. L.; TAVARES, M. B.; LEMOS, J. A.; GRISWOLD, A. R.; RIBEIRO, J. L.; BALAN, A.; GUIMARÃES, K. S.; CAI, S.; BURNE, R. A.; FERREIRA, L. C. S. et al. The oligopeptide (opp) gene cluster of Streptococcus mutans: identification, prevalence, and characterization. **Oral microbiology and immunology**, v. 22, p. 277-84, 2007.

NIEBISCH, A.; KABUS, A.; SCHULTZ, C.; WEIL, B.; BOTT, M. Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. **The Journal of biological chemistry**, v. 281, p. 12300-7, 2006.

NOZAKI CN, FARIA MAR & MACHADO TMM. Extirpação cirúrgica dos abscessos da LC em caprinos. **Arq. Inst. Biol.** 67: pp. 187-9. 2000

NYSTRÖM, T.; NEIDHARDT, F. C. Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in *Escherichia coli*. **Molecular microbiology**, v. 6, n. 21, p. 3187-3198, 1992.

OKUNO, D.; IINO, R.; NOJI, H. Rotation and structure of FoF1-ATP synthase. **Journal of biochemistry**, v. 149, p. 655-64, 2011.

OLIVER, H. F.; ORSI, R. H.; PONNALA, L.; KEICH, U.; WANG, W.; SUN, Q.; CARTINHO, S. W.; FILIATRAULT, M. J.; WIEDMANN, M.; BOOR, K. J. Deep RNA sequencing of *L. monocytogenes* reveals overlapping and extensive stationary phase and sigma B-dependent transcriptomes, including multiple highly transcribed noncoding RNAs. **BMC genomics**, v. 10, p. 641, 2009.

OLIVEIRA DM. Sequenciamento do genoma do *corynebacterium pseudotuberculosis* e as implicações no diagnóstico e controle da LC. **II simpósio de higiene e sanidade animal** .2007

OLSEN, I.; REITAN, L. J.; HOLSTAD, G.; WIKER, H. G. Alkyl hydroperoxide reductases C and D are major antigens constitutively expressed by *Mycobacterium avium* subsp. paratuberculosis. **Infection and immunity**, v. 68, p. 801-8, 2000.

OLSON, M. E.; CERI, H.; MORCK, D. W.; BURET, A. G.; READ, R. R. Biofilm bacteria: formation and comparative susceptibility to antibiotics. **Canadian journal of veterinary research = Revue canadienne de recherche vétérinaire**, v. 66, p. 86-92, 2002.

ORAM DM, JACOBSON AD, HOLMES RK. Transcription of the contiguous sigB, dtxR, and galE genes in *Corynebacterium diphtheriae*: evidence for multiple transcripts and regulation by environmental factors. **J Bacteriol.** 188(8):2959-73.2006

PACHECO, L.G.C. **Caracterização do exoproteoma basal e variante de *Corynebacterium pseudotuberculosis* e a relação com a virulência reduzida de uma linhagem mutante para um fator sigma alternativo**. Tese (Doutorado em Bioquímica e Imunologia) –Universidade Federal de Minas Gerais, BH, MG,(2010), pp:1-193

PACHECO L.G.C., PENA R.R., CASTRO T.L.P., DORELLA F.A., BAHIA R.C., CARMINATI R., FROTA M.N.L., OLIVEIRA S.C., MEYER R., ALVES F.S.F., MIYOSHI A., AZEVEDO V. Multiplex PCR Assay for Identification of *Corynebacterium pseudotuberculosis* from Pure Cultures and for Rapid Detection of this Pathogen in Clinical Samples. **J. Med. Microbiol.**, 2007.

PACHECO, L.; SLADE, S.; SEYFFERT, N.; SANTOS, A.; CASTRO, T.; SILVA, W.; SANTOS, A.; SANTOS, S.; FARIAS, L.; CARVALHO, M. et al. A combined approach for

comparative exoproteome analysis of *Corynebacterium pseudotuberculosis*. **BMC Microbiology**, v. 11, n. 1, p. 12, 2011.

PADAN E, VENTURI M, GERCHMAN Y, DOVER N. Na(+)/H(+) antiporters. **Biochim Biophys Acta**. 1505(1):144-57. 2001

PAITHOONRANGSARID K, SHOUMSKAYA MA, KANESAKI Y, SATOH S, TABATA S, LOS DA, ZINCHENKO VV, HAYASHI H, TANTICHAROEN M, SUZUKI I, MURATA N. Five histidine kinases perceive osmotic stress and regulate distinct sets of genes in *Synechocystis*. **J Biol Chem**. 279 (51):53078-86. 2004

PAO, S. S.; PAULSEN, I. A. N. T.; SAIER, M. H. Major Facilitator Superfamily. **Microbiology**, v. 62, p. 1-34, 1998.

PARSOT, C.; MEKALANOS, J. J. Expression of ToxR, the transcriptional activator of the virulence factors in *Vibrio cholerae*, is modulated by the heat shock response. **Proceedings of the National Academy of Sciences**, v. 87, n. 24, p. 9898-9902, December 1, 1990

PASSALACQUA, K. D.; VARADARAJAN, A.; ONDOV, B. D.; OKOU, D. T.; ZWICK, M. E.; BERGMAN, N. H. Structure and complexity of a bacterial transcriptome. **Journal of bacteriology**, v. 191, p. 3203-11, 2009.

PATON MW, ROSE IR, HART RA, SUTHERLAND SS, MERCY AR, ELLIS TM, DHALIWAL JA. New infection with *corynebacterium pseudotuberculosis* reduces wool production. **Aust. Vet. J.** 71: pp. 47-9. 1994

PATON MW, WALKER SB, ROSE IR & WATT GF. Prevalence of caseous lymphadenitis and usage of caseous lymphadenitis vaccines in sheep flocks. **Aust. Vet. J.** 81: pp. 91-5.2003

PEARCE, S. R.; MIMMACK, M. L.; GALLAGHER, M. P.; GILEADI, U.; HYDE, S. C.; HIGGINS, C. F. Membrane topology of the integral membrane components, OppB and OppC, of the oligopeptide permease of *Salmonella typhimurium*. **Molecular microbiology**, v. 6, n. 1, p. 47-57, 1992.

PERERA, I. C.; GROVE, A. MarR homologs with urate-binding signature. **Protein Science**, v. 20, n. 3, p. 621-629, 2011.

PERKINS, T. T.; KINGSLEY, R. A.; FOOKES, M. C.; GARDNER, P. P.; JAMES, K. D.; YU, L.; ASSEFA, S. A.; HE, M.; CROUCHER, N. J.; PICKARD, D. J. et al. A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*. **PLoS genetics**, v. 5, p. e1000569, 2009.

PETIT, C.; BROWN, J.; INGRAHAM, K.; BRYANT, A.; HOLMES, D. Lipid modification of prelipoproteins is dispensable for growth in vitro but essential for virulence in *Streptococcus pneumoniae*. **FEMS Microbiol Lett**, v. 200(2), p. 229-33, 2001a.

R.D. FINN, J. MISTRY, J. TATE, P. COGGILL, A. HEGER, J.E. POLLINGTON, O.L. GAVIN, P. GUNESEKARAN, G. CERIC, K. FORSLUND, L. HOLM, E.L. SONNHAMMER, S.R. EDDY, A. BATEMAN. Pfam protein families database. *Nucleic Acids Research* (2010).

PIERSON, D. E.; FALKOW, S. Nonpathogenic isolates of *Yersinia enterocolitica* do not contain functional inv-homologous sequences. **Infect. Immun.**, v. 58, n. 4, p. 1059-1064, April 1, 1990 1990.

PORTEVIN, D.; DE SOUSA-D'AURIA, C.; HOUSSIN, C.; GRIMALDI, C.; CHAMI, M.; DAFFÉ, M.; GUILHOT, C. A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. **Proceedings of the National Academy of Sciences**, v. 101, n. 1, p. 314-319, 2004.

PORTEVIN, D.; DE SOUSA-D'AURIA, C.; MONTROZIER, H.; HOUSSIN, C.; STELLA, A.; LANÉELLE, M.-A.; BARDOU, F.; GUILHOT, C.; DAFFÉ, M. The acyl-AMP ligase FadD32 and AccD4-containing acyl-CoA carboxylase are required for the synthesis of mycolic acids and essential for mycobacterial growth: identification of the carboxylation product and determination of the acyl-CoA carboxylase componen. **The Journal of biological chemistry**, v. 280, p. 8862-74, 2005.

PRICE CW, FAWCETT P, CÉRÉMONIE H, SU N, MURPHY CK, YOUNGMAN P. Genome-wide analysis of the general stress response in *Bacillus subtilis*. **Mol Microbiol.** 41(4):757-74. 2001.

PUGH DG. **Caseous lymphadenitis**. In: *Sheep & Goat Medicine* Saunders: pp. 207-8. 2002

QUIGLEY, G.; TEETER, M.; RICH, A. Structural analysis of spermine and magnesium ion binding to yeast phenylalanine transfer RNA. **Proc Natl Acad Sci USA**, v. 75, p. 64-8, 1978.

QUIVEY, R. J.; KUHNERT, W.; HAHN, K. Adaptation of oral streptococci to low pH. **Adv Microb Physiol**, v. 42, p. 239-74, 2000.

RAMAN S, SONG T, PUYANG X, BARDAROV S, JACOBS WR JR, HUSSON RN. The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. **J Bacteriol.** 183(20):6119-25. 2001

RAMOS, J. L.; MARTI, M.; MOLINA-HENARES, A. J.; TERA, W.; BRENNAN, R.; TOBES, R. The TetR Family of Transcriptional Repressors. **Society**, v. 69, p. 326-356, 2005.

RAMOS, R. T.; CARNEIRO, A. R.; BAUMBACH, J.; AZEVEDO, V.; SCHNEIDER, M. P.; SILVA, A. Analysis of quality raw data of second generation sequencers with Quality Assessment Software. **BMC Res Notes**, v. 4, p. 130, 2011.

RASMUSSEN, S.; NIELSEN, H. B.; JARMER, H. The transcriptionally active regions in the genome of *Bacillus subtilis*. **Molecular microbiology**, v. 73, p. 1043-57, 2009.

RATHMAN, M.; SJAASTAD, M.; FALKOW, S. Acidification of phagosomes containing *Salmonella typhimurium* in murine macrophages. **Infect. Immun.**, v. 64, n. 7, p. 2765-2773, July 1, 1996 1996.

RATLEDGE C, DOVER LG Iron metabolism in pathogenic bacteria. **Annu Rev Microbiol.** 54:881-941.2000

RIBEIRO, O.; SILVA, J.; OLIVEIRA, S.; MEYER, R.; FERNANDES, G. Preliminary results on a living vaccine against caseous lymphadenitis. **Pesquisa Agropecuaria Brasileira**, v. 26, p. 461 - 465, 1991.

RIGALI, S.; SCHLICHT, M.; HOSKISSON, P.; NOTHAFT, H.; MERZBACHER, M.; JORIS, B.; TITGEMEYER, F. Extending the classification of bacterial transcription factors beyond the helix-turn-helix motif as an alternative approach to discover new cis/trans relationships. **Nucleic Acids Res**, v. 32, p. 3418 - 3426, 2004.

ROBERTS, A.; TRAPNELL, C.; DONAGHEY, J.; RINN, J. L.; PACHTER, L. Improving RNA-Seq expression estimates by correcting for fragment bias. **Genome biology**, v. 12, p. R22, 2011.

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

ROHMER, L.; HOCQUET, D.; MILLER, S. I. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. **Trends in microbiology**, p. 1-8, 2011.

RUSSELL, N. J., EVANS, R. I., TER STEEG, P. F., HELLEMONS, J., VERHEUL, A. ABEE, T. Membranes as a target for stress adaptation. **Int J Food Microbiol** 28, 255–261, 1995

SANCAR, A. DNA Excision Repair. **Annual review of biochemistry**, v. 65, n. 1, p. 43-81, 1996.

SANTOS, L.P. Caprinos e ovinos: informações importantes. Natal:SEBRAE/RN. 90p. 2006.

SASINDRAN, S. J.; SAIKOLAPPAN, S.; DHANDAYUTHAPANI, S. Methionine sulfoxide reductases and virulence of bacterial pathogens. **Future microbiology**, v. 2, p. 619-30, 2007.

SATORHELYI, P. Microarray Analyse der pH Stressantwort von *Listeria monocytogenes* und *Corynebacterium glutamicum*. 2004.

SAVIOLA, B.; WOOLWINE, S. C.; BISHAI, W. R. Isolation of Acid-Inducible Genes of *Mycobacterium tuberculosis* with the Use of Recombinase-Based In Vivo Expression Technology. **Infect. Immun.**, v. 71, n. 3, p. 1379-1388, March 1, 2003 2003.

SCHELL, M. A. Molecular Biology of the LysR family of transcriptional regulators **Annu. Rev. Microbiol**, v. 47, p. 597-626, 1993.

SCHMITT, M.P. Utilization of host iron sources by *Corynebacterium diphtheriae*: identification of a gene whose product is homologous to eukaryotic heme oxygenases and is required for acquisition of iron from heme and hemoglobin. **J Bacteriol** 179: 838-45. 1997

SCHMITT, M. P.; HOLMES, R. K. Iron-dependent regulation of diphtheria toxin and siderophore expression by the cloned *Corynebacterium diphtheriae* repressor gene dtxR in *C. diphtheriae* C7 strains. **Infection and immunity**, v. 59, p. 1899-904, 1991.

SCHNEEWIND, O.; FOWLER, A.; FAULL, K. F. Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. **Science**, v. 268(5207), p. 103-6. 1995.

SCHÖNBRUNNER, E. R.; SCHMID, F. X. Peptidyl-prolyl cis-trans isomerase improves the efficiency of protein disulfide isomerase as a catalyst of protein folding. **Proceedings of the National Academy of Sciences of the United States of America**, v. 89, p. 4510-3, 1992.

SCHUMANN, W. Thermosensors in eubacteria: role and evolution. **J.Biosci**, v. 32, p. 549-57, 2007.

SCHUMANN W, EHRLICH SD e OGASAWARA N. Functional analysis of bacterial genes. John Wiley e Sons, Ltd., Chichester, United Kingdom .2001

SCHWARZER, N.; NÖST, R.; SEYBOLD, J.; PARIDA, S. K.; FUHRMANN, O.; KRÜLL, M.; SCHMIDT, R.; NEWTON, R.; HIPPENSTIEL, S.; DOMANN, E. et al. Two distinct phospholipases C of *Listeria monocytogenes* induce ceramide generation, nuclear factor-kappa B activation, and E-selectin expression in human endothelial cells. **Journal of immunology (Baltimore, Md. : 1950)**, v. 161, p. 3010-8, 1998.

SCOTT, J. R.; BARNETT, T. C. Surface proteins of gram-positive bacteria and how they get there. **Annual review of microbiology**, v. 60, p. 397-423, 2006.

SELIM SA. Oedematous skin disease of buffalo in egypt. **J. Vet. Med. B Infect. Dis. Vet. Public Health** 48: pp. 241-58. 2001

SENDER, E.; JOHNSON, G.D.; KRAWETZ, S.A. Local and global factors affecting RNA sequencing analysis. **Anal Biochem**. 2011

SERGANOV, A. The long and the short of *riboswitches*. **Current opinion in structural biology**, v. 19, p. 251-9, 2009.

SERVANT, P.; MAZODIER, P. Negative regulation of the heat shock response in *Streptomyces*. **Archives of Microbiology**, v. 176, n. 4, p. 237-242, 2001.

SEYFFERT N, GUIMARÃES AS, PACHECO LGC, PORTELA RW, BASTOS BL, DORELLA FA, HEINEMANN MB, LAGE AP, GOUVEIA AMG, MEYER R, MIYOSHI A e AZEVEDO V. High seroprevalence of caseous lymphadenitis in brazilian goat herds revealed by *Corynebacterium pseudotuberculosis* secreted proteins-based elisa. **Res. Vet. Sci**. 88: pp. 50-5. 2010

SHAO C, ZHANG Q, TANG W, QU W, ZHOU Y, SUN Y, YU H, JIA J. The changes of proteomes components of *Helicobacter pylori* in response to acid stress without urea. **J Microbiol**. 46(3):331-7. 2008

SHARMA, C. M.; HOFFMANN, S.; DARFEUILLE, F.; REIGNIER, J.; FINDEISS, S.; SITTKA, A.; CHABAS, S.; REICHE, K.; HACKERMÜLLER, J.; REINHARDT, R. et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. **Nature**, v. 464, p. 250-5, 2010.

SHIN, N.-R.; LEE, D. Y.; YOO, H. S. Identification of quorum sensing-related regulons in *Vibrio vulnificus* by two-dimensional gel electrophoresis and differentially displayed reverse transcriptase PCR. **FEMS immunology and medical microbiology**, v. 50, p. 94-103, 2007.

SHPIGEL NY, ELAD D, YERUHAM I, WINKLER M, SARAN A. An outbreak of *Corynebacterium pseudotuberculosis* infection in an Israeli dairy herd. **Vet Rec.** 133(4):89-94. 1993

SLEATOR RD e HILL C. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. **FEMS Microbiol. Rev.** 26: pp. 49-71. 2002

SOARES SC. **Validação de um método computacional para identificação, caracterização e comparação in silico de ilhas de patogenicidade no gênero *Corynebacterium* e aplicação no genoma de duas linhagens de *Corynebacterium pseudotuberculosis*.** Dissertação (Mestrado em Genética)-Universidade Federal de Minas Gerais, BH, MG.

SONG, T.; SONG, S.-E.; RAMAN, S.; ANAYA, M.; HUSSON, R. N. Critical role of a single position in the -35 element for promoter recognition by *Mycobacterium tuberculosis* SigE and SigH. **Journal of bacteriology**, v. 190, p. 2227-30, 2008.

SOREK, R.; COSSART, P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. **Nature reviews. Genetics**, v. 11, p. 9-16, 2010.

SOTO, G. E.; HULTGREN, S. J. Bacterial adhesins: common themes and variations in architecture and assembly. **Journal of bacteriology**, v. 181, p. 1059-71, 1999.

SOUTOURINA, J.; PLATEAU, P.; BLANQUET, S. Metabolism of d-Aminoacyl-tRNAs in *Escherichia coli* and *Saccharomyces cerevisiae* Cells. **Journal of Biological Chemistry**, v. 275, n. 42, p. 32535-32542, 2000

SOUZA, B.M. **Construção de um mutante para o gene *sigH* codificador do fator sigma alternativo σ^H e análise do papel desse fator na resposta de *Corynebacterium pseudotuberculosis* a diferentes condições de estresse ambiental.** Dissertação (Mestrado em Genética)- Universidade Federal de Minas Gerais (2011), pp: 1-114.

SPOHN, G.; SCARLATO, V. The autoregulatory HspR repressor protein governs chaperone gene transcription in *Helicobacter pylori*. **Molecular microbiology**, v. 34, p. 663-74, 1999.

SRITHARAN, M. Iron and bacterial virulence. **Indian journal of medical microbiology**, v. 24, p. 163-4, 2006.

STALLINGS, C. L.; STEPHANOU, N. C.; CHU, L.; HOCHSCHILD, A.; NICKELS, B. E.; GLICKMAN, M. S. CarD Is an Essential Regulator of rRNA Transcription Required for *Mycobacterium tuberculosis* Persistence. **Cell**, v. 138, n. 1, p. 146-159, 2009.

STANFORD K, BROGDEN KA, MCCLELLAND LA, KOZUB GC, AUDIBERT F. The incidence of caseous lymphadenitis in alberta sheep and assessment of impact by vaccination with commercial and experimental vaccines. **Can. J. Vet. Res.** 62: pp. 38-43. 1998

- STEIL, L.; HOFFMANN, T.; BUDDE, I.; VO, U.; BREMER, E. Genome-Wide Transcriptional Profiling Analysis of Adaptation of *Bacillus subtilis* to High Salinity. **Society**, v. 185, p. 6358-6370, 2003.
- STEWART, G. R.; SNEWIN, V. A.; WALZL, G.; HUSSELL, T.; TORMAY, P.; O'GAORA, P.; GOYAL, M.; BETTS, J.; BROWN, I. N.; YOUNG, D. B. Overexpression of heat-shock proteins reduces survival of *Mycobacterium tuberculosis* in the chronic phase of infection. **Nat Med**, v. 7, n. 6, p. 732-737, 2001.
- STEWART, G. R.; WERNISCH, L.; STABLER, R.; MANGAN, J. A.; HINDS, J.; LAING, K. G.; YOUNG, D. B.; BUTCHER, P. D. Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. **Microbiology**, v. 148, n. 10, p. 3129-3138, 2002.
- STOCK AM, ROBINSON VL, GOUDREAU PN. Two-component signal transduction. **Annu Rev Biochem** 69:183-215. 2000
- SONGER JG. Bacterial phospholipases and their role in virulence. **Trends Microbiol.** 5: pp. 156-61. 1997
- SONGER JG, LIBBY SJ, IANDOLO JJ & CUEVAS WA. Cloning and expression of the phospholipase D gene from *Corynebacterium pseudotuberculosis* in escherichia coli. **Infect. Immun.** 58: pp. 131-6. 1990
- SOREK R, COSSART P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. **Nat. Rev. Genet.** 11: pp. 9-16. 2010
- STOREY, A. D. A. J. D. qvalue. **Journal of the Royal Statistical Society**, v. 64, p. 479-498, 2002.
- STORZ, G.; IMLAY, J. oxidative stress. **current opinion in microbiology**, v. 2, p. 188-194, 1999.
- STURGILL-KOSZYCKI, S.; SCHLESINGER, P.; CHAKRABORTY, P.; HADDIX, P.; COLLINS, H.; FOK, A.; ALLEN, R.; GLUCK, S.; HEUSER, J.; RUSSELL, D. Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. **Science**, v. 263, n. 5147, p. 678-681, 1994.
- SUN R, CONVERSE PJ, KO C, TYAGI S, MORRISON NE, BISHAI WR. *Mycobacterium tuberculosis* ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set. **Mol Microbiol.** 52(1):25-38. 2004
- SZEGEDY, M. A.; MAGUIRE, M. E. The CorA Mg(2+) transport protein of *Salmonella typhimurium*. Mutagenesis of conserved residues in the second membrane domain. **The Journal of biological chemistry**, v. 274, p. 36973-9, 1999.
- TAKAYAMA, K.; KJELLEBERG, S. The role of RNA stability during bacterial stress responses and starvation. **Environmental Microbiology**, v. 2, n. 4, p. 355-365, 2000.

TASHJIAN JJ & CAMPBELL SG. Interaction between caprine macrophages and *Corynebacterium pseudotuberculosis*: an electron microscopic study. **Am. J. Vet. Res.** 44: pp. 690-3. 1983

TAYLOR WE, STRAUS DB, GROSSMAN AD, BURTON ZF, GROSS CA, BURGESS RR. Transcription from a heat-inducible promoter causes heat shock regulation of the sigma subunit of *E. coli* RNA polymerase. **Cell.** 38(2):371-81. 1984

TENG, X.; XIAO, H. Perspectives of DNA microarray and next-generation DNA sequencing technologies. **Science in China. Series C, Life sciences / Chinese Academy of Sciences**, v. 52, p. 7-16, 2009.

TESTA, B.; KRÄMER, S. D. The Biochemistry of Drug Metabolism – An Introduction. **Chemistry e Biodiversity**, v. 3, n. 10, p. 1053-1101, 2006.

TIAN G; WK., M. Mutational analysis of the arginine repressor of *Escherichia coli*. **Mol Microbiol**, v. 13, p. 599-608, 1994.

TJALSMA, H.; KONTINEN, V. P.; PRÁGAI, Z.; WU, H.; MEIMA, R.; VENEMA, G.; BRON, S.; SARVAS, M.; VAN DIJL, J. M. The Role of Lipoprotein Processing by Signal Peptidase II in the Gram-positive Eubacterium *Bacillus subtilis*. **Journal of Biological Chemistry**, v. 274, n. 3, p. 1698-1707, 1999.

TOKISHITA, S.-I.; MIZUNO, T. Transmembrane signal transduction by the *Escherichia coli* osmotic sensor, EnvZ: intermolecular complementation of transmembrane signalling. **Molecular microbiology**, v. 13, n. 3, p. 435-444, 1994.

TOKUNAGA, M.; TOKUNAGA, H.; WU, H. C. Post-translational modification and processing of *Escherichia coli* prolipoprotein in vitro. **Proceedings of the National Academy of Sciences**, v. 79, n. 7, p. 2255-2259, 1982.

TOLEDO-ARANA, A.; DUSSURGET, O.; NIKITAS, G.; SESTO, N.; GUET-REVILLET, H.; BALESTRINO, D.; LOH, E.; GRIPENLAND, J.; TIENSUU, T.; VAITKEVICIUS, K. et al. The *Listeria* transcriptional landscape from saprophytism to virulence. **Nature**, v. 459, p. 950-6, 2009.

TOLEDO-ARANA, A.; REPOILA, F.; COSSART, P. Small noncoding RNAs controlling pathogenesis. **Current opinion in microbiology**, v. 10, n. 2, p. 182-188, 2007.

TOMBOLA, F.; MORBIATO, L.; DEL GIUDICE, G.; RAPPUOLI, R.; ZORATTI, M.; PAPINI, E. The *Helicobacter pylori* VacA toxin is a urea permease that promotes urea diffusion across epithelia. **The Journal of Clinical Investigation**, v. 108, n. 6, p. 929-937, 2001.

TON-THAT, H.; SCHNEEWIND, O. Assembly of pili on the surface of *Corynebacterium diphtheriae*. **Molecular Microbiology**, v. 50, p. 1429-1438, 2003.

TON-THAT, H.;SCHNEEWIND, O. Assembly of pili in Gram-positive bacteria. **Trends in microbiology**, v. 12, p. 228-34, 2004.

TRIPATHI, L. P.; SOWDHAMINI, R. Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. **BMC genomics**, v. 9, p. 549, 2008.

TRIVEDI, O. A.; ARORA, P.; SRIDHARAN, V.; TICKOO, R. Enzymic activation and transfer of fatty acids as acyl-adenylates in mycobacteria biosynthetic intermediates are transferred from one enzymatic. **Nature**, v. 428, p. 4-6, 2004.

UEHARA, T.; PARK, J. T. The N-Acetyl- D -Glucosamine Kinase of *Escherichia coli* and Its Role in Murein Recycling. **Society**, v. 186, p. 7273-7279, 2004.

UNANIAN, M.; FELICIANO, S. A.; PANT , K. Abscesses and caseous lymphadenitis in goats in tropical semi-arid north-east Brazil. **Trop Anim Health Prod**, v. 17, p. 57-62, 1985.

VANEECHOUTTE, M. et al. Evaluation of the applicability of amplified rDNA-restriction analysis (ARDRA) to identification of species of the genus *Corynebacterium*. **Res.Microbiol.**, n. 146, p. 633-641, 1995.

VAN VLIET, A.; BERESWILL, S.; KUSTERS, J. Ion Metabolism and Transport. In: MOBLEY HLT, M. G., HAZELL SL (Ed.). **Helicobacter pylori: Physiology and Genetics**. Washington (DC): ASM Press, cap. 17, 2001.

VAN VLIET, A. H. M. Next generation sequencing of microbial transcriptomes: challenges and opportunities. **FEMS microbiology letters**, v. 302, p. 1-7, 2010.

VIJAY, K.; BRODY, M. S.; FREDLUND, E.; PRICE, C. W. A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the σ B transcription factor of *Bacillus subtilis*. **Molecular microbiology**, v. 35, n. 1, p. 180-188, 2000.

VINDAL, V.; RANJAN, S.; RANAJAN, A. In silico analysis and characterization of GntR family of regulators from *Mycobacterium tuberculosis*. **Tuberculosis**, v. 87, p. 242 - 247, 2007.

VRIESEMA, A. J. M.; DANKERT, J.; ZAAT, S. A. J. A Shift from Oral to Blood pH Is a Stimulus for Adaptive Gene Expression of *Streptococcus gordonii* CH1 and Induces Protection against Oxidative Stress and Enhanced Bacterial Growth by Expression of msrA. **Infect. Immun.**, v. 68, n. 3, p. 1061-1068, 2000.

WAISBERG M, LOBO FP, CERQUEIRA GC, PASSOS LKJ, CARVALHO OS, FRANCO GR & EL-SAYED NM. **Microarray analysis of gene expression induced by sexual contact in *Schistosoma mansoni***. **BMC Genomics** 8: p. 181. 2007

WALKER, J.; JACKSON, H. J.; EGGLETON, D. G.; MEEUSEN, E. N.; WILSON, M. J.; BRANDON, M. R. Identification of a novel antigen from *Corynebacterium pseudotuberculosis* that protects sheep against caseous lymphadenitis. **Infection and immunity**, v. 62, p. 2562-7, 1994.

WANG, L.; FENG, Z.; WANG, X.; WANG, X.; ZHANG, X. DEGseq : an R package for identifying differentially expressed genes from RNA-seq data. **Genome**, p. 1-3, 2009.

- WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq : a revolutionary tool for transcriptomics. **October**, v. 10, p. 57-63, 2009.
- WANG QP, KAGUNI JM. A novel sigma factor is involved in expression of the rpoH gene of *Escherichia coli*. **J Bacteriol** 171(8):4248-53.1989
- WEBB, S. A. R.; KAHLER, C. M. Bench-to-bedside review: Bacterial virulence and subversion of host defences. **Critical care (London, England)**, v. 12, p. 234, 2008.
- WEEKS, D. L.; ESKANDARI, S.; SCOTT, D. R.; SACHS, G. A H⁺-Gated Urea Channel: The Link Between *Helicobacter pylori* Urease and Gastric Colonization. **Science**, v. 287, n. 5452, p. 482-485, January 21, 2000 2000.
- WEEKS, D. L.; SACHS, G. Sites of pH regulation of the urea channel of *Helicobacter pylori*. **Molecular microbiology**, v. 40, n. 6, p. 1249-1259, 2001.
- WEERASINGHE, J. P.; DONG, T.; SCHERTZBERG, M. R.; KIRCHHOF, M. G.; SUN, Y.; SCHELLHORN, H. E. Stationary phase expression of the arginine biosynthetic operon argCBH in *Escherichia coli*. **BMC microbiology**, v. 6, p. 14, 2006.
- WEIHOFEN, W. A.; BERGER, M.; CHEN, H.; SAENGER, W.; HINDERLICH, S. Structures of human N-Acetylglucosamine kinase in two complexes with N-Acetylglucosamine and with ADP/glucose: insights into substrate specificity and regulation. **Journal of molecular biology**, v. 364, p. 388-99, 2006.
- WEINBERG ED. Iron and infection. **Microbiol Rev.** 42(1):45-66. 1978
- WESCHE, A. M.; GURTLER, J. B.; MARKS, B. P.; RYSER, E. T. Stress, sublethal injury, resuscitation, and virulence of bacterial foodborne pathogens. **Journal of food protection**, v. 72, p. 1121-38, 2009.
- WEST, N. P.; CHOW, F. M. E.; RANDALL, E. J.; WU, J.; CHEN, J.; RIBEIRO, J. M. C.; BRITTON, W. J. Cutinase-like proteins of *Mycobacterium tuberculosis*: characterization of their variable enzymatic functions and active site identification. **The FASEB journal : official publication of the Federation of American Societies for Experimental Biology**, v. 23, p. 1694-704, 2009.
- WILLIAMSON L.H. Caseous lymphadenitis in mall ruminants, Vet. Clin. North Am. **Food Anim. Pract.** 17, p. 359–37, 2001.
- WILSON, M. J.; BRANDON, M. R.; WALKER, J. Molecular and biochemical characterization of a protective 40-kilodalton antigen from *Corynebacterium pseudotuberculosis*. **Infection and immunity**, v. 63, p. 206-11, 1995.
- WU QL, KONG D, LAM K, HUSSON RN. A mycobacterial extracytoplasmic function sigma factor involved in survival following stress. **J Bacteriol.** 179(9):2922-9(1997)
- WURTZEL, O.; SAPRA, R.; CHEN, F.; ZHU, Y.; SIMMONS, B. A.; SOREK, R. A single-base resolution map of an archaeal transcriptome. **Genome research**, v. 20, p. 133-41, 2010.

- YANG, H.; ZHENG, G.; PENG, X.; QIANG, B.; YUAN, J. D-Amino acids and D-Tyr-tRNA^{Tyr} deacylase: stereospecificity of the translation machine revisited. **FEBS Letters**, v. 552, p. 95-98, 2003.
- YERUHAM I, BRAVERMAN Y, SHPIGEL NY, CHIZOV-GINZBURG A, SARAN A, WINKLER M. Mastitis in dairy cattle caused by *corynebacterium pseudotuberculosis* and the feasibility of transmission by houseflies. i. **Vet Q** 18: pp. 87-9. 1996
- YERUHAM I, FRIEDMAN S, PERL S, ELAD D, BERKOVICH Y, KALGARD Y. A herd level analysis of a *corynebacterium pseudotuberculosis* outbreak in a dairy cattle herd. **Vet. Dermatol.** 15: pp. 315-20. 2004
- YUAN, Y, CRANE, DD, SIMPSON, RM, ZHU, YQ, HICKEY, MJ, SHERMAN, DR, BARRY, CE. The 16-kDa alpha-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. **Proc Natl Acad Sci USA** 95: 9578–9583.1998
- YUAN, Y.; LEE, R. E.; BESRA, G. S.; BELISLE, J. T.; BARRY, C. E. Identification of a gene involved in the biosynthesis of cyclopropanated mycolic acids in *Mycobacterium tuberculosis*. **Proceedings of the National Academy of Sciences**, v. 92, n. 14, p. 6630-6634, 1995.
- YURA, T.; NAKAHIGASHI, K. Regulation of the heat-shock response. **Current opinion in microbiology**, v. 2, n. 2, p. 153-158, 1999.
- ZAHRT, T. C.; DERETIC, V. An Essential Two-Component Signal Transduction System in *Mycobacterium tuberculosis*. **J. Bacteriol.**, v. 182, n. 13, p. 3832-3838, 2000.
- ZAIM, J.; KIERZEK, A. M. The structure of full-length LysR-type transcriptional regulators. Modeling of the full-length OxyR transcription factor dimer. **Nucleic acids research**, v. 31, n. 5, p. 1444-1454, 2003.
- ZHOU, D., YANG, R. Global analysis of gene transcription regulation in prokaryotes. **Cell. Mol. Life Sci.** 63: pp. 2260-90. 2006

10. ANEXOS

Genômica : Produção Científica

Artigos completos publicados

Anexo 1

Ruiz JC, D'Afonseca V, Silva A, Ali A, Pinto AC, Santos AR, Rocha AA, Lopes DO, Dorella FA, Pacheco LG, Costa MP, Turk MZ, Seyffert N, Moraes PM, Soares SC, Almeida SS, Castro TL, Abreu VA, Trost E, Baumbach J, Tauch A, Schneider MP, McCulloch J, Cerdeira LT, Ramos RT, Zerlotini A, Dominittini A, Resende DM, Coser EM, Oliveira LM, Pedrosa AL, Vieira CU, Guimarães CT, Bartholomeu DC, Oliveira DM, Santos FR, Rabelo EM, Lobo FP, Franco GR, Costa AF, Castro IM, Dias SR, Ferro JA, Ortega JM, Paiva LV, Goulart LR, Almeida JF, Ferro MI, Carneiro NP, Falcão PR, Grynberg P, Teixeira SM, Brommonschenkel S, Oliveira SC, Meyer R, Moore RJ, Miyoshi A, Oliveira GC, Azevedo V. **Evidence for reductive genome evolution and lateral acquisition of virulence functions in two *Corynebacterium pseudotuberculosis* strains.** PLoS One. 2011 6(4):e18551.

OPEN ACCESS Freely available online



Evidence for Reductive Genome Evolution and Lateral Acquisition of Virulence Functions in Two *Corynebacterium pseudotuberculosis* Strains

Jerônimo C. Ruiz¹⁰, Vivian D'Afonseca²⁰, Artur Silva³, Amjad Ali², Anne C. Pinto², Anderson R. Santos², Aryanne A. M. C. Rocha², Débora O. Lopes⁴, Fernanda A. Dorella², Luis G. C. Pacheco²⁰, Marcília P. Costa⁵, Meritxell Z. Turk², Núbia Seyffert², Pablo M. R. O. Moraes², Siomar C. Soares², Sintia S. Almeida², Thiago L. P. Castro², Vinicius A. C. Abreu², Eva Trost⁶, Jan Baumbach⁷, Andreas Tauch⁶, Maria Paula C. Schneider², John McCulloch⁸, Louise T. Cerdeira³, Rommel T. J. Ramos³, Adhemar Zerlotini¹, Anderson Dominittini¹, Daniela M. Resende^{1,8}, Elisângela M. Coser¹, Luciana M. Oliveira⁹, André L. Pedrosa^{8,10}, Carlos U. Vieira¹¹, Cláudia T. Guimarães¹², Daniela C. Bartholomeu¹³, Diana M. Oliveira⁵, Fabricio R. Santos², Elida Mara Rabelo¹⁴, Francisco P. Lobo¹³, Glória R. Franco¹³, Ana Flávia Costa², Ieso M. Castro¹⁵, Sílvia Regina Costa Dias¹⁴, Jesus A. Ferro¹⁶, José Miguel Ortega¹³, Luciano V. Paiva¹⁷, Luiz R. Goulart¹¹, Juliana Franco Almeida¹¹, Maria Inês T. Ferro¹⁶, Newton P. Carneiro¹², Paula R. K. Falcão¹⁸, Priscila Grynberg¹³, Santuza M. R. Teixeira¹³, Sérgio Brommonschenkel¹⁹, Sérgio C. Oliveira¹³, Roberto Meyer²⁰, Robert J. Moore²¹, Anderson Miyoshi², Guilherme C. Oliveira^{1,22}, Vasco Azevedo^{23,9}

¹ Research Center René Rachou, Oswaldo Cruz Foundation, Belo Horizonte, Minas Gerais, Brazil, ² Department of General Biology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ³ Department of Genetics, Federal University of Pará, Belém, Pará, Brazil, ⁴ Health Sciences Center, Federal University of São João del Rei, Ouro Preto, Minas Gerais, Brazil, ⁵ Department of Veterinary Medicine, State University of Ceará, Fortaleza, Ceará, Brazil, ⁶ Department of Genetics, University of Bielefeld, Bielefeld, Nordrhein Westfalen, Germany, ⁷ Department of Computer Science, Max Planck Institut für Informatik, Saarbrücken, Saarland, Germany, ⁸ Department of Pharmaceutical Sciences, Federal University of Ouro Preto, Ouro Preto, Minas Gerais, Brazil, ⁹ Department of Physics, Federal University of Ouro Preto, Ouro Preto, Minas Gerais, Brazil, ¹⁰ Department of Biological Sciences, Federal University of Triângulo Mineiro, Uberaba, Minas Gerais, Brazil, ¹¹ Department of Genetics and Biochemistry, Federal University of Uberlândia, Uberlândia, Minas Gerais, Brazil, ¹² Brazilian Agricultural Research Corporation (EMBRAPA), Sete Lagoas, Minas Gerais, Brazil, ¹³ Department of Biochemistry and Immunology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ¹⁴ Department of Parasitology, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil, ¹⁵ Department of Pharmacy, Federal University of Ouro Preto, Ouro Preto, Minas Gerais, Brazil, ¹⁶ Department of Technology, State University of São Paulo, Jaboticabal, São Paulo, Brazil, ¹⁷ Department of Chemistry, Federal University of Lavras, Lavras, Minas Gerais, Brazil, ¹⁸ Brazilian Agricultural Research Corporation (EMBRAPA), Campinas, São Paulo, Brazil, ¹⁹ Department of Plant Pathology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil, ²⁰ Department of Biotechnology Sciences, Federal University of Bahia, Salvador, Bahia, Brazil, ²¹ CSIRO Livestock Industries, Australia, ²² Center of Excellence in Bioinformatics, National Institute of Science and Technology, Research Center René Rachou, Oswaldo Cruz Foundation, Belo Horizonte, Minas Gerais, Brazil

Abstract

Background: *Corynebacterium pseudotuberculosis*, a Gram-positive, facultative intracellular pathogen, is the etiologic agent of the disease known as caseous lymphadenitis (CL). CL mainly affects small ruminants, such as goats and sheep; it also causes infections in humans, though rarely. This species is distributed worldwide, but it has the most serious economic impact in Oceania, Africa and South America. Although *C. pseudotuberculosis* causes major health and productivity problems for livestock, little is known about the molecular basis of its pathogenicity.

Methodology and Findings: We characterized two *C. pseudotuberculosis* genomes (Cp1002, isolated from goats; and Cp2231, isolated from sheep). Analysis of the predicted genomes showed high similarity in genomic architecture, gene content and genetic order. When *C. pseudotuberculosis* was compared with other *Corynebacterium* species, it became evident that this pathogenic species has lost numerous genes, resulting in one of the smallest genomes in the genus. Other differences that could be part of the adaptation to pathogenicity include a lower GC content, of about 52%, and a reduced gene repertoire. The *C. pseudotuberculosis* genome also includes seven putative pathogenicity islands, which contain several classical virulence factors, including genes for fimbrial subunits, adhesion factors, iron uptake and secreted toxins. Additionally, all of the virulence factors in the islands have characteristics that indicate horizontal transfer.

Anexo 2

Silva A, Schneider MP, Cerdeira L, Barbosa MS, Ramos RT, Carneiro AR, Santos R, Lima M, D'Afonseca V, Almeida SS, Santos AR, Soares SC, Pinto AC, Ali A, Dorella FA, Rocha F, de Abreu VA, Trost E, Tauch A, Shpigel N, Miyoshi A, Azevedo V. **Complete genome sequence of *Corynebacterium pseudotuberculosis* I19, a strain isolated from a cow in Israel with bovine mastitis.** J Bacteriol. (2011) 193(1):323-4

JOURNAL OF BACTERIOLOGY, Jan. 2011, p. 323–324
0021-9193/11/\$12.00 doi:10.1128/JB.01211-10
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

Vol. 193, No. 1

Complete Genome Sequence of *Corynebacterium pseudotuberculosis* I19, a Strain Isolated from a Cow in Israel with Bovine Mastitis[▽]

Artur Silva,¹ Maria Paula C. Schneider,¹ Louise Cerdeira,¹ Maria Silvanira Barbosa,¹ Rommel Thiago J. Ramos,¹ Adriana R. Carneiro,¹ Rodrigo Santos,¹ Marília Lima,¹ Vivian D'Afonseca,² Sintia S. Almeida,² Anderson R. Santos,² Siomar C. Soares,² Anne C. Pinto,² Amjad Ali,² Fernanda A. Dorella,² Flavia Rocha,² Vinicius Augusto Carvalho de Abreu,² Eva Trost,³ Andreas Tauch,³ Nahum Shpigel,⁴ Anderson Miyoshi,² and Vasco Azevedo^{2*}

Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brazil¹; Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil²; CeBiTec, Universität Bielefeld, 33594 Bielefeld, Germany³; and the Koret School of Veterinary Medicine, Hebrew University of Jerusalem, P.O. Box 12, Rehovot 76100, Israel⁴

Received 8 October 2010/Accepted 19 October 2010

This work reports the completion and annotation of the genome sequence of *Corynebacterium pseudotuberculosis* I19, isolated from an Israeli dairy cow with severe clinical mastitis. To present the whole-genome sequence, a *de novo* assembly approach using 33 million short (25-bp) mate-paired SOLiD reads only was applied. Furthermore, the automatic, functional, and manual annotations were attained with the use of several algorithms in a multistep process.

Anexo 3

Louise Cerdeira, Anne Pinto, Maria Paula Schneider, Sintia Almeida, Anderson Santos, Eudes Guilherme Barbosa, Amjad Ali, Maria Barbosa, Adriana Carneiro, Rommel Thiago Ramos, Rodrigo Santos, Debmalya Barh, Neha Barve, Vasudeo Zambare, Silvia Belchior, Luis Carlos Guimarães, Siomar Soares, Fernanda Dorella, Flavia Rocha, Vinicius Augusto Abreu, Andreas Tauch, Eva Trost, Anderson Miyoshi, Vasco Azevedo, and Artur Silva. **Whole genome sequence of *Corynebacterium pseudotuberculosis* PAT10 of strain isolated from sheep in Patagonia, Argentine.** J.Bacteriol

Whole-Genome Sequence of *Corynebacterium pseudotuberculosis* PAT10 Strain Isolated from Sheep in Patagonia, Argentina

Louise Teixeira Cerdeira,¹ Anne Cybelle Pinto,² Maria Paula Cruz Schneider,¹ Sintia Silva de Almeida,² Anderson Rodrigues dos Santos,² Eudes Guilherme Vieira Barbosa,² Amjad Ali,² Maria Silvanira Barbosa,¹ Adriana Ribeiro Carneiro,¹ Rommel Thiago Jucá Ramos,¹ Rodrigo Santos de Oliveira,¹ Debmalya Barh,⁴ Neha Barde,⁴ Vasudeo Zambare,^{3,5} Sílvia Estevão Belchior,² Luis Carlos Guimarães,² Siomar de Castro Soares,² Fernanda Alves Dorella,² Flavia Souza Rocha,² Vinicius Augusto Carvalho de Abreu,² Andreas Tauch,⁶ Eva Trost,⁶ Anderson Miyoshi,² Vasco Azevedo,² and Artur Silva^{1*}

Instituto de Ciências Biológicas Universidade Federal do Pará, Belém, Brazil¹; Instituto de Ciências Biológicas Universidade Federal de Minas Gerais Belo Horizonte, Brazil²; Universidad Nacional de la Patagonia San Juan Bosco, Chubut, Argentina³; Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIoAB), Howrah, Purba Medinipur, West Bengal, India⁴; Centre for Bioprocessing Research and Development (CBRD), South Dakota School of Mines and Technology, Rapid City, South Dakota⁵; and CeBITec, University of Bielefeld, 33594 Bielefeld, Germany⁶

Received 20 August 2011/Accepted 26 August 2011

In this work, we report the complete genome sequence of a *Corynebacterium pseudotuberculosis* PAT10 isolate, collected from a lung abscess in an Argentine sheep in Patagonia, whose pathogen also required an investigation of its pathogenesis. Thus, the analysis of the genome sequence offers a means to better understanding of the molecular and genetic basis of virulence of this bacterium.

The incidence of caseous lymphadenitis (CLA) is high in many regions of the world, resulting in huge and significant economic losses in agribusinesses, as it is responsible for a decrease in wool production and carcass quality (3). The disease is endemic in flocks in the provinces of Chubut and Santa Cruz in the Southern Patagonia region of Argentina, thereby leading to an outrageous prevalence rate of about 70% within individual flocks in Patagonia (4). Diseases caused by *Corynebacterium pseudotuberculosis* present in various clinical forms, and sheep and goats are affected by CLA (1). Analysis of the genome sequence improves our understanding of the molecular and genetic basis of the virulence of the bacterium (2). We hereby report the whole-genome sequence of *C. pseudotuberculosis* PAT10 as determined using the SOLiD platform. In total, we generated 27,858,221 mate-paired short reads (25 bp) of usable sequences (286-fold coverage). Furthermore, a hybrid *de novo* assembly approach was applied using 16,885,903 short (25-bp) mate-paired SOLiD filtered reads; that strategy allowed close gaps without a bench work time cost (2). The automatic and manual annotations were done using several algorithms in a multistep process.

For structural annotation, the following software was used: Prodigal (gene predictor); RNAMmer (rRNA predictor) (5);

(8). Manual annotation was performed using Artemis software (7). Identification and confirmation of pseudogenes in the genome were carried out using CLCBio Workbench 4.0.2 software. Manual analysis was performed based on the Phred quality of each base combined with analysis of the depth of coverage of the frameshift region. That analysis allowed the identification of false-positive pseudogene results. The genome of the PAT10 strain consists of a 2,555,323-bp circular chromosome, and the average G+C content of the chromosome is 52.19%. The genome was predicted to contain 2,079 coding genes (CDS), four rRNA operons, 49 rRNAs, and 61 tRNAs.

The characterization of the PAT10 genome should identify and unravel the mechanisms of virulence of this pathogen through comparative analyses performed with other sequenced genomes of the genus and the same species, thereby allowing the development of new diagnostic kits and vaccines.

Nucleotide sequence accession number. The genome sequence obtained in this study has been deposited in the GenBank database under accession number CP002924 (chromosome).

Downloaded from http://j.asm.org/ on November 21, 2011 by guest

Artigos aceitos para publicação

Anexo 4

Louise Cerdeira, Maria Paula Schneider, Anne Pinto, Sintia Almeida, Anderson Santos, Eudes Guilherme Barbosa, Amjad Ali, Flavia Aburjaile, Vinicius Augusto Abreu, Luis Carlos Guimarães, Siomar Soares, Fernanda Dorella, Flavia Rocha, Erick Bol, Pablo Sá, Thiago Lopes, Maria Barbosa, Adriana Carneiro, Rommel Thiago Ramos, Nilson Antônio Coimbra, Alex Ranieri Lima, Debmalya Barh, Neha Jain, Sandeep Tiwari, Rathiram Raja, Vasudeo Zambare, Preetam Ghosh, Eva Trost, Andreas Tauch, Anderson Miyoshi, Vasco Azevedo, and Artur Silva. **Complete Genome Sequence of *Corynebacterium pseudotuberculosis* strain - CIP 52.97- isolated from Horse in Kenya.** J.Bacteriol

Re: JB06293-11 (Complete Genome Sequence of *Corynebacterium pseudotuberculosis* strain - CIP 52.97- isolated from Horse in Kenya)

Dear Prof. Silva:

Your manuscript has been accepted, and I am forwarding it to the ASM Journals Department for publication. For your reference, ASM Journals' address is given below. Before it can be scheduled for publication, your manuscript must be checked by the ASM production editor to make sure that all elements meet the technical requirements for publication. Stefanie Kowalski, the production editor for Journal of Bacteriology, will contact you if anything needs to be revised before copyediting and production can begin. Otherwise, you will be notified when your proofs are ready to be viewed.

1 Complete Genome Sequence of *Corynebacterium pseudotuberculosis* strain CIP
2 52.97, isolated from Horse in Kenya.

3 Louise Teixeira Cerdeira¹, Maria Paula Cruz Schneider¹, Anne Cybelle Pinto², Sinitia
4 Silva de Almeida², Anderson Rodrigues dos Santos², Eudes Guilherme Vieira Barbosa²,
5 Amjad Ali³, Maria Silvanira Barbosa¹, Adriana Ribeiro Carneiro¹, Rommel Thiago Jucá
6 Ramos¹, Rodrigo Oliveira dos Santos¹, Nilson Coimbra¹, Alex Ranieri Jerônimo Lima¹,
7 Erick Bol¹, Debmalya Barh⁴, Neha Barve⁴, Vasudeo Zambare^{4,5}, Siomar de Castro
8 Soares², Anderson Miyoshi⁶, Vasco Azevedo², Artur Silva^{1*}

9

10 ¹Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Brazil.

11 ²Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo
12 Horizonte, Brazil.

13 ³Universidad Nacional de la Patagonia San Juan Bosco, Chubut, Argentina.

14 ⁴Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and
15 Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal, India.

16 ⁵Center for Bioprocessing Research and Development (CBRD), South Dakota School of
17 Mines and Technology, South Dakota, USA.

18

23 In this work, we have reported the whole genome sequence of *Corynebacterium*
24 *pseudotuberculosis* CIP 52.97 (Collection Institut Pasteur biovar equi), isolated in
25 1952 from lymphangitis in Kenya horse, which brings great losses to
26 agrobusiness, thereby obtaining the genome will allowed detect targets important
27 in post genomic studies, aiming to minimize the problems caused by this
28 microorganism.

29

30 *Corynebacterium pseudotuberculosis*, intracellular pathogen, brings great harm to the
31 creators of goats, sheep, horses, cattle, since the infected animals have wounds in the
32 skin and or internal organs causing damage to the skin of the animal and even the
33 condemnation of the flesh (Dorella et al., 2006). This bacterium is classified by two
34 biovars: type I, ovis, nitrate negative, type II, equi, nitrate positive, and the described
35 strain in this paper, CIP5297, belongs to biovar type II. In horses, three forms of the
36 disease are described: external abscesses, ulcerative lymphangitis, and so visceral,
37 affecting internal organs (Aleman et al., 1996). Lymphangitis (also known as 'chest
38 abscess', 'pigeon breast', 'pigeon fever') is one of the most common and economically
39 important infectious diseases of horses (Doherr et al., 1998). For knowledge of the
40 molecular and genetic basis of virulence of this bacterium it was necessary sequencing

Capítulo de Livro – aceito para publicação

Anexo 5

Abreu, V. A. C., Almeida, S., Santos, A. R., Soares, S. C., Ali, A., Pinto, A. C., Magalhães. A.A.C., Magalhães. A., Barbosa, E. G. V., Ramos, R. T. J., Cerdeira, L. T., Carneiro, A. R., Schneider, M. P. C., Silva, A., Miyoshi, A., Azevedo, V. **Whole Genome Annotation: In Silico Analysis. In: Bioinformatics / Book 1, 2011, v.1, p. 1-16.**

Chapter Number

2 Whole Genome Annotation: In Silico Analysis

3 Vasco Azevedo et al.*
4 Federal University of Minas Gerais (UFMG) and Federal University of Pará (UFPA),
5 Brazil

6 1 Introduction

7 After a genome is assembled, the next step is genomic annotation, which can generate data
8 that will allow various types of research of the model organism. Complete DNA sequences
9 of the organism are then mapped in areas pertinent to the research objectives. In this
10 chapter, we explore relevant ongoing research on genes and consider the gene as a basic
11 mapping unit. Gene prediction is the first hurdle we come across to begin the extensive and
12 intensive work demonstrated in first item, which deals with assembly of the genome. Gene
13 prediction can be made with computational techniques for recognizing gene sequences,
14 including stop codons and the initial portions of nucleotide sequences; it involves empirical
15 rules concerning minimum coding sequences (CDS's) and is limited due to overlapping
16 sequences coding forward and reverse.
17 Finishing gene prediction step by a computer initiates the functional annotation stage.
18 Functional annotation, item 3, can be done initially by computer, using similarity in
19 sequence alignment. However, no software is capable of generating a functional
20 annotation without many false positive results, since conserved protein domains with
21 varied functions make gene sequence alignment difficult. In this case, after automatic
22 annotation, the predicted genes need to be revised manually. In manual curation, item 4,
23 an expert can more accurately locate frameshifts in the DNA strand. Depending on the
24 number of errors found, genomic annotation may be postponed, requiring a return to the
25 previous stage of genome assembly. In manual curation, the principal contributions are
26 usually correction of the start codon position, gene name, gene product and, finally,
27 identification of frameshifts.
28 When functional annotation is completed, the genome should subsequently be submitted. It
29 occurs after the assembly and annotation steps: making the data generated available in
30 public-access databanks. Submission is a pre-requisite for publication in scientific journals.
31 Another advantage of genome publication in public-access sites is that it permits use of
32 various genome analysis tools. For example, searches for genomic plasticity, pangenomic
33 study, exported antigens and evaluation of innate and adaptive immune responses. The
34 pangenome approach, item 5, concept of species can be used as a filter for targeting
35 candidates for vaccines, diagnostic kit and drug development. For drug development the

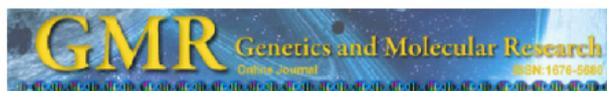
* Vinícius Abreu, Sinitia Almeida, Anderson Santos, Siomar Soares, Amjad Ali, Anne Pinto, Assymé Magalhães, Eudes Barbosa, Rommel Ramos, Louise Cerdeira, Adriana Carneiro, Paula Schneider, Artur Silva and Anderson Miyoshi
Federal University of Minas Gerais (UFMG) and Federal University of Pará (UFPA), Brazil

Transcriptômica: Produção Científica

Artigo completo publicado

Anexo 6

Pinto AC, Melo-Barbosa HP, Miyoshi A, Silva A, Azevedo V.
Application of RNA-seq to reveal the transcript profile in bacteria. Genet Mol Res.
10(3):1707-18. 2011



Review

Application of RNA-seq to reveal the transcript profile in bacteria

A. C. Pinto¹, H.P. Melo-Barbosa², A. Miyoshi¹, A. Silva¹ and V. Azevedo¹

¹Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

²Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brasil

Corresponding author: V. Azevedo
E-mail: vasco@icb.ufmg.br

Genet. Mol. Res. 10 (3): 1707-1718 (2011)
Received July 7, 2011
Accepted August 4, 2011
Published August 16, 2011
DOI <http://dx.doi.org/10.4238/vol10-3gmr1554>

ABSTRACT. The large number of microbial genomes deposited in databanks has opened the door for in-depth studies of organisms, including post-genomics investigations. Thanks to new generation sequencing technology, these studies have made advances that have led to extraordinary discoveries in bacterial transcriptomics. In this review, we describe bacterial RNA sequencing studies that use these new techniques. We also examined the advantages and biases of these new generation technologies; advances in bioinformatics make it possible to overcome the biases, providing interesting and surprising results.

Key words: RNA-seq; Transcriptome; Bacteria

Genetics and Molecular Research 10(3): 1707-1718 (2011)

©FUNPEC-EP www.funpec.org.br

Artigo submetido

Anexo 7

Anne Cybelle Pinto, Rommel T J Ramos, Wanderson Marques Silva, Flávia Souza Rocha, Silvanira Barbosa, Maria P C Schneider, Artur Silva, Vasco Azevedo. **Core Stimulon and new transcripts discovery of *Corynebacterium pseudotuberculosis* strain 1002 based on *ab initio* methodologies.** Submetido a revista Integrative Biology (2011).

Cite this: DOI: 10.1039/c6cc00000x

www.rsc.org/xxxxx

ARTICLETYPE

The core stimulon and new transcripts of *Corynebacterium pseudotuberculosis* strain 1002 identified using *ab initio* methodologies

Anne Cybelle Finto^{1,1}, Rommel T J Ramos^{1,1}, Wanderson Marques Silva², Flávia Souza Rocha², Silvanira Barbosa², Anderson Miyoshi², Maria P C Schneider², Artur Silva², Vasco Azevedo²

¹ These authors contributed equally to this work

Received (in XXX, XXX) XXX XXXXXXXX XXXX, Accepted (in XXX, XXX) XXXXXXXX XXXX

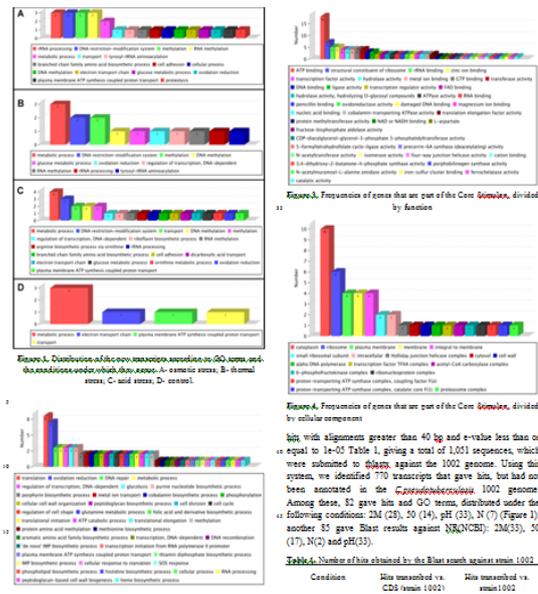
DOI: 10.1039/xxxxxxx

Abstract
Corynebacterium pseudotuberculosis is a bacterium that causes diseases such as caseous lymphadenitis in small ruminants, resulting in large-scale economic losses for agribusiness worldwide. Consequently, it has been the focus of various studies, including analysis of its transcriptomic profile. Identification of the transcripts that appear under conditions that simulate the environment encountered by this bacterial species in the host is of great importance for the effort to discover new targets for the production of more efficient vaccines. We sequenced the cDNA of *Corynebacterium pseudotuberculosis* strain 1002, using the SOLiD V3 system, under the following conditions: aerobic stress (2M), acidity (pH), heat shock (50°C) and control condition (C). To identify the transcripts shared among the conditions, we integrate this information with the results from BLAST and BLAST2GO, we developed the software *CoposStimulon* (CS), which allows the user to individually characterize the genes in terms of their participation in biological processes, their function and cellular location. The data generated by this software was also used to identify new transcripts in the genome. Considering metabolic processes, we observed that three genes are represented in the core stimulon and seven genes in the control; this supports the hypothesis that the organism's strategy for surviving in a hostile environment is through a reduction in growth. The DNA repair process, the ATP catabolic process, and cell adhesion are controlled by genes that contribute to bacterial cell maintenance under stress conditions; these could be involved in their pathogenicity. The methodology for the identification of transcripts obtained by *ab initio* assembly and that are shared among the conditions allowed us to select candidates for vaccine studies, principally through the implementation of the software *CoposStimulon*, which allows a search for genes associated with pathogenicity of the organism, besides permitting identification of new transcripts that had not been previously annotated; this will allow updating of the annotation of the genome in future studies.

Introduction

Corynebacterium pseudotuberculosis is a pathogenic bacterium that is Gram-positive, belongs to the CMDE group, and causes diseases such as caseous lymphadenitis, which mainly affects goats and sheep, resulting in economic losses for goat and sheep farmers. Unfortunately, few vaccine targets have been identified, which makes development of proper control measures difficult. Using the new sequencing technologies, it has become possible to obtain complete number of reads from the transcriptome to regress genes involved in ribosomal translation and biogenesis¹⁻⁴ these cause the cessation or reduction in growth that is generally observed in response to stress, this being an important strategy for physiological cellular adjustment to a novel environmental condition.⁴

In this study, we used SOLiD^{5,6} technology to investigate the transcriptomic profile of *Corynebacterium pseudotuberculosis* strain 1002 (*C. pseudotuberculosis* 1002) under conditions that simulate the host environment: BHI supplemented with sodium chloride (2M), with hydrochloric acid (pH), to simulate the acidity of



Gene	Controle	2M	50°C	pH	fold-change 2M	fold-change 50°C	fold-change pH
Cp1002_2031	5,95	25,69	28,32	82,6	4,32	4,76	13,88
Cp1002_1418	0	33,58	41,12	12,85	33,58	41,12	12,85
Cp1002_0730	5,18	92,54	35,64	57,41	17,86	6,88	11,08
dps	1493,78	1261,38	2146,15	12269,57	0,84	1,44	8,21
Cp1002_0531	79,85	171,34	143,3	619,03	2,15	1,79	7,75
acnA	126,56	168,36	186,21	853,33	1,33	1,47	6,74
dtd	7,51	51,88	33,09	43,44	6,91	4,41	5,78
grpE	552,84	499,31	2015,54	3080,09	0,90	3,65	5,57
Cp1002_1434	26,12	72,42	72,69	140,42	2,77	2,78	5,38
Cp1002_0025	26,14	155,39	56,92	131,15	5,94	2,18	5,02
Cp1002_0758	5,63	52,24	32,73	26,5	9,28	5,81	4,71
phuC	4,94	20,12	20,91	23,11	4,07	4,23	4,68
Cp1002_0968	17,01	132,23	53,98	78,73	7,77	3,17	4,63
Cp1002_1870	22,12	189,53	47,98	83,79	8,57	2,17	3,79
ruvC	64,54	210,84	135,53	237,46	3,27	2,10	3,68
sigH	320	598,88	595,45	1175,3	1,87	1,86	3,67
Cp1002_0810	81,66	278,73	153,53	282,78	3,41	1,88	3,46
Cp1002_0268	1402,52	5777,26	3987,89	4608,57	4,12	2,84	3,29
panC	62,66	141,26	112,74	203,35	2,25	1,80	3,25
Cp1002_0526	34,26	59,64	63,97	105,63	1,74	1,87	3,08
Cp1002_2001	433,46	705,59	623,04	1300,59	1,63	1,44	3,00
Cp1002_0898	164,66	223,47	241,67	493,2	1,36	1,47	3,00
Cp1002_1471	510,7	803,73	986,98	1529,58	1,57	1,93	3,00
Cp1002_0729	13,76	63,85	37,23	40,31	4,64	2,71	2,93
Cp1002_0321	45,66	98,32	99,01	131,81	2,15	2,17	2,89

Cp1002_0746	22,68	94,39	61,06	65,44	4,16	2,69	2,89
Cp1002_0090	28,06	109,86	53,42	79,77	3,92	1,90	2,84
mraZ	1363,52	2031,87	2056,4	3866,44	1,49	1,51	2,84
tatB	207,22	258,37	281,39	579,98	1,25	1,36	2,80
hemK	110,34	344,43	257,99	308,49	3,12	2,34	2,80
rpsT	1220,77	3711,33	1973,65	3349,54	3,04	1,62	2,74
Cp1002_0077	35,77	61,23	67,41	97,39	1,71	1,88	2,72
Cp1002_0176	39,45	247,57	69,53	106,93	6,28	1,76	2,71
Cp1002_1659	147,57	171,07	189,1	398,13	1,16	1,28	2,70
Cp1002_1787	113,94	266,05	178,52	299,92	2,34	1,57	2,63
trpC2	51	102,95	74,24	133,98	2,02	1,46	2,63
Cp1002_1933	57,32	123,62	104,96	150,39	2,16	1,83	2,62
Cp1002_0893	90,56	228,93	204,02	236,72	2,53	2,25	2,61
Cp1002_0662	146,13	165,9	202,43	381,33	1,14	1,39	2,61
Cp1002_1469	138,19	275,07	187,9	358,77	1,99	1,36	2,60
Cp1002_0457	11,51	56,57	29,49	29,41	4,91	2,56	2,56
dnaJ	344,19	335,54	1284,9	862,13	0,97	3,73	2,50
Cp1002_0757	11,63	43,2	27,06	28,84	3,71	2,33	2,48
yihR	369,06	77,83	197,1	913,74	0,21	0,53	2,48
ruvA	157,45	266,57	224,84	377,54	1,69	1,43	2,40
infA	1856,82	3786,52	4196,89	4440,6	2,04	2,26	2,39
Cp1002_0677	212,12	412,64	408,79	500,02	1,95	1,93	2,36
gltA	149,12	67,48	286,98	349,49	0,45	1,92	2,34
degV	56,99	123,34	103,23	132,53	2,16	1,81	2,33
araC	385,15	47,52	385,51	892,41	0,12	1,00	2,32
Cp1002_1947	14,08	39,18	44,67	32,58	2,78	3,17	2,31
Cp1002_0797	116,6	226,21	136,34	269,52	1,94	1,17	2,31
Cp1002_1547	34,71	84,78	72,36	78,67	2,44	2,08	2,27
Cp1002_1134	11,88	29,85	30,84	26,78	2,51	2,60	2,25
corA	20,84	81,21	55,91	45,37	3,90	2,68	2,18
rplN	1015,03	2954,68	2903,66	2203,3	2,91	2,86	2,17
Cp1002_1751	89,97	370,06	248,36	195,07	4,11	2,76	2,17
sdrD	109,02	268,89	150,96	235,64	2,47	1,38	2,16
hemH1	218	528	360,64	469,2	2,42	1,65	2,15
rplX	1045,59	2426,86	2865,91	2248,53	2,32	2,74	2,15
rplW	1040,76	1428,88	1499,56	2217,63	1,37	1,44	2,13
ylxR	831,84	2792,06	1886,89	1772,15	3,36	2,27	2,13
Cp1002_1054	48,86	96,44	132,86	103,82	1,97	2,72	2,12
prfA	535,25	1146,18	1135,89	1137,17	2,14	2,12	2,12
hemG	101,73	174,69	203,96	215,79	1,72	2,00	2,12
Cp1002_0517	48,89	115,03	85,85	103,18	2,35	1,76	2,11
Cp1002_0625	21,72	46,59	42,24	45,16	2,15	1,94	2,08
Cp1002_0199	198,62	245,6	326,54	410,04	1,24	1,64	2,06
rpsA	5354,38	12128,18	10000,34	10914,67	2,27	1,87	2,04
Cp1002_0544	55,26	178,64	106,62	112,56	3,23	1,93	2,04
mntA	51,2	6,88	72,79	103,8	0,13	1,42	2,03
adk	526,45	1267,51	1028,11	1061,91	2,41	1,95	2,02
rplE	1559,84	3672,83	4478,32	3123,29	2,35	2,87	2,00
Cp1002_0958	273,85	800,68	492	546,4	2,92	1,80	2,00
rpsJ	1627,86	2147,94	2903,16	3244,41	1,32	1,78	1,99
rfbE	36,98	72,74	63,01	72,66	1,97	1,70	1,96
sufR	422,72	439,17	437,38	828,61	1,04	1,03	1,96
rpmE	1080,64	2474,01	2126,14	2110,88	2,29	1,97	1,95
obgE	110,88	210,91	165,9	216,37	1,90	1,50	1,95
rpsK	650,59	1035,49	1033,49	1251,66	1,59	1,59	1,92

lysA	172,07	473,33	330,66	330,31	2,75	1,92	1,92
dhh	242,87	495,24	339,18	459,39	2,04	1,40	1,89
aroA	58,07	98,45	92,12	109,66	1,70	1,59	1,89
secA	605,39	539,64	647,48	1142,94	0,89	1,07	1,89
Cp1002_1335	520,69	647,35	716,4	981,55	1,24	1,38	1,89
Cp1002_1634	81,93	107,6	128,18	154,44	1,31	1,56	1,89
Cp1002_0107	23,72	64,53	51,43	44,69	2,72	2,17	1,88
regX3	115,76	219,54	147,24	217,96	1,90	1,27	1,88
topA	101,76	177,85	151,73	191,53	1,75	1,49	1,88
udg	39,12	113,22	75,71	73,47	2,89	1,94	1,88
trkH	74,22	166,22	111,27	139,11	2,24	1,50	1,87
Cp1002_1320	154,1	185,32	243,7	288,49	1,20	1,58	1,87
pac2	373,13	400,14	394,05	697,97	1,07	1,06	1,87
Cp1002_1631	62,57	106,28	98,01	117,01	1,70	1,57	1,87
Cp1002_1385	45,82	88,65	77,54	85,43	1,93	1,69	1,86
rplI	1803,85	2806,36	3338,78	3342,79	1,56	1,85	1,85
Cp1002_0715	172,44	422,53	397,61	319,14	2,45	2,31	1,85
Cp1002_1085	409,43	807,09	689,69	752,42	1,97	1,68	1,84
deaD	1542,6	5746,82	3088,32	2817,14	3,73	2,00	1,83
Cp1002_1522	83,34	422,64	212,75	151,45	5,07	2,55	1,82
rpmA	1625,04	2989,94	2893,79	2947,96	1,84	1,78	1,81
Cp1002_1201	144,3	205,81	210,41	261,25	1,43	1,46	1,81
rpsN	1889,39	2685,69	2726,3	3401,42	1,42	1,44	1,80
purH	229,77	444,82	337,86	413,61	1,94	1,47	1,80
mntD	28,23	54,56	54,02	50,65	1,93	1,91	1,79
trkA	43,09	131,67	77,99	77,24	3,06	1,81	1,79
rpsP	2286,48	5139,29	1943,9	4097,07	2,25	0,85	1,79
mtrB	194,76	364,33	249,3	348,85	1,87	1,28	1,79
rpsR	2801,54	3230,93	4532,31	5006,57	1,15	1,62	1,79
infC	951,18	1638,69	1643,5	1696,85	1,72	1,73	1,78
Cp1002_0509	176,3	185,54	206,83	313,75	1,05	1,17	1,78
Cp1002_0598	271,83	536,03	360,33	481,09	1,97	1,33	1,77
ureG	147,83	220,17	205,02	261,59	1,49	1,39	1,77
yidC1	356,68	654,87	588,98	630,67	1,84	1,65	1,77
rplM	2085,64	2936,31	3807,15	3684,13	1,41	1,83	1,77
nagD	357,95	944,6	474,52	632,02	2,64	1,33	1,77
Cp1002_1823	76,88	98,62	106,42	134,54	1,28	1,38	1,75
Cp1002_1294	76,65	125,64	111,26	133,8	1,64	1,45	1,75
Cp1002_0485	98,39	156,72	166,48	171,28	1,59	1,69	1,74
secE	199,78	773,52	385,59	346,69	3,87	1,93	1,74
scpA	72,45	145,24	118,8	125,24	2,00	1,64	1,73
malC	286,71	307,42	291,35	493,81	1,07	1,02	1,72
Cp1002_1387	176,17	293,84	225,5	302,17	1,67	1,28	1,72
Cp1002_1048	27,16	73,56	48,3	46,52	2,71	1,78	1,71
Cp1002_1753	70,32	112,64	93,56	120,34	1,60	1,33	1,71
Cp1002_1262	592,68	865,38	794,31	1014,2	1,46	1,34	1,71
menA	78,93	119,68	128,82	134,96	1,52	1,63	1,71
Cp1002_0815	61,02	142,44	100,42	104,27	2,33	1,65	1,71
Cp1002_0988	99,82	711,84	189,65	170,41	7,13	1,90	1,71
Cp1002_1031	137,47	170,96	192,49	233,94	1,24	1,40	1,70
ftn	77,18	1135,74	262,5	131,34	14,72	3,40	1,70
Cp1002_1253	132,61	341,16	188,96	225,63	2,57	1,42	1,70
Cp1002_0411	400,39	377,01	532,03	679,46	0,94	1,33	1,70
rsmB	116,79	191,16	188,49	198,12	1,64	1,61	1,70
infB	678,71	1365,16	1072,88	1147,98	2,01	1,58	1,69

phoH	295,28	791,73	484,22	498,77	2,68	1,64	1,69
Cp1002_0572	862,34	388,55	463,13	1455,96	0,45	0,54	1,69
nrdH	286,16	367,68	492,09	482,18	1,28	1,72	1,69
gidB	362,12	555,5	479,78	608,94	1,53	1,32	1,68
rbfA	539,43	702	645,41	904,68	1,30	1,20	1,68
pknB	69,58	108,41	101,3	116,09	1,56	1,46	1,67
purS	124,76	154,76	165,53	206,94	1,24	1,33	1,66
rpmD	3099,57	5560,9	5810,02	5128,6	1,79	1,87	1,65
spoU1	410,9	534,53	444,95	677,39	1,30	1,08	1,65
Cp1002_1183	87,18	194,62	148,64	142,31	2,23	1,70	1,63
purF	385,92	911,12	508,33	629,65	2,36	1,32	1,63
rplD	2110,17	2035,16	2037,98	3402,55	0,96	0,97	1,61
helY	78,79	115,45	124,46	126,69	1,47	1,58	1,61
rho	483,41	1231,16	936,4	773,86	2,55	1,94	1,60
mhpD	333,6	411,23	451,65	531,73	1,23	1,35	1,59
Cp1002_0204	342,04	592,48	467,36	543,09	1,73	1,37	1,59
Cp1002_0767	217,93	454,65	333,81	345,49	2,09	1,53	1,59
Cp1002_1300	138,9	190,38	180,54	219,5	1,37	1,30	1,58
Cp1002_0253	29,78	86,8	60,03	47,06	2,91	2,02	1,58
ruvB	130,4	225,23	166,74	206,01	1,73	1,28	1,58
rpsI	2050,16	2828,82	2946,62	3221,93	1,38	1,44	1,57
pyrR	465,11	536,51	472,87	730,34	1,15	1,02	1,57
Cp1002_0124	383,25	512,84	527,05	596,93	1,34	1,38	1,56
Cp1002_1398	345,01	784,55	411,07	536,52	2,27	1,19	1,56
Cp1002_0478	126,68	141,93	166,57	196,93	1,12	1,31	1,55
rpsM	1038,39	1402,83	1303,41	1607,06	1,35	1,26	1,55
ribD	284,82	282,42	303,26	440,25	0,99	1,06	1,55
Cp1002_0796	332,62	343,25	375,37	513,97	1,03	1,13	1,55
Cp1002_0594	45,37	158,98	97,8	70,08	3,50	2,16	1,54
rpsS	1129,78	1193,07	1155,61	1741,46	1,06	1,02	1,54
nifU	400,44	690,6	667,86	615,43	1,72	1,67	1,54
Cp1002_1655	421,39	535,7	419,81	646,27	1,27	1,00	1,53
aspS	690,18	726,64	692,49	1057,81	1,05	1,00	1,53
betP	624,47	1486,68	1294,26	956,39	2,38	2,07	1,53
trxA1	124,34	140,75	165,26	189,95	1,13	1,33	1,53
Cp1002_0825	545,26	513,43	552,89	830,34	0,94	1,01	1,52
rplT	587,76	922,27	794,83	894,98	1,57	1,35	1,52
Cp1002_0431	34,53	69,34	53,5	52,55	2,01	1,55	1,52
impA	40,33	94,27	61,77	61,36	2,34	1,53	1,52
Cp1002_0811	310,56	479,62	457,24	472,22	1,54	1,47	1,52
cspB	260,43	391,19	276,88	394,24	1,50	1,06	1,51
relA	275,78	562,94	335,85	417,44	2,04	1,22	1,51
Cp1002_1789	46,28	93,61	70,84	70,05	2,02	1,53	1,51
terC	75,78	200,79	137,39	114	2,65	1,81	1,50
proC	80,35	212,23	120,39	120,85	2,64	1,50	1,50
Cp1002_1152	75,46	101,76	106,48	112,93	1,35	1,41	1,50
Cp1002_1764	67,91	196,14	97,72	101,57	2,89	1,44	1,50
ubiA	204,77	300,97	272,9	305,42	1,47	1,33	1,49
nudF	218,09	379,49	361,45	325,2	1,74	1,66	1,49
Cp1002_2069	451,07	189,18	810,15	670,96	0,42	1,80	1,49
Cp1002_0876	303,63	402,73	328,08	450,38	1,33	1,08	1,48
Cp1002_0245	1045,37	1239	1364,72	1547,27	1,19	1,31	1,48
Cp1002_0543	81,16	186,32	140,2	119,97	2,30	1,73	1,48
Cp1002_1591	195,43	481,11	264,38	288,17	2,46	1,35	1,47
dnaX	121,9	177,37	150,37	179,1	1,46	1,23	1,47

hemC	152,14	231,12	180,28	222,97	1,52	1,18	1,47
secY	227,55	698,29	379,91	332,97	3,07	1,67	1,46
nusA	167,3	362,85	262,8	244,26	2,17	1,57	1,46
gltX1	436,47	449,61	422,52	634,91	1,03	0,97	1,45
mscL	895,75	939,19	1765,33	1299,22	1,05	1,97	1,45
Cp1002_1336	1339,89	827,72	1263,49	1939,8	0,62	0,94	1,45
murA	193,16	275,28	259,24	279,59	1,43	1,34	1,45
tlyA	199,35	570,81	299,08	288,32	2,86	1,50	1,45
rplB	1186,96	1102,59	1103,67	1716,26	0,93	0,93	1,45
Cp1002_1569	57,13	107,38	89,74	82,48	1,88	1,57	1,44
dapC	115,84	160,15	182,45	167,13	1,38	1,58	1,44
uppP	47,21	197,48	105,68	68,09	4,18	2,24	1,44
Cp1002_0684	65,99	167,6	131,96	94,59	2,54	2,00	1,43
Cp1002_0073	63,03	148,38	84,4	90,23	2,35	1,34	1,43
ribH	264,16	513,79	374,11	378,12	1,94	1,42	1,43
mshB	229,45	294,84	279,74	328,18	1,28	1,22	1,43
rpoA	1232,26	1599,06	1690,99	1761,29	1,30	1,37	1,43
cytA	72,93	312,46	169,19	103,71	4,28	2,32	1,42
eno	6812,66	6386,01	7230,09	9655,67	0,94	1,06	1,42
rpsF	3644,45	5388,76	5507,51	5158,59	1,48	1,51	1,42
Cp1002_1614	130,71	460,29	182,52	184,98	3,52	1,40	1,42
Cp1002_0596	70,11	99,73	94,87	99,18	1,42	1,35	1,41
cwlM	189,76	301,1	294,69	267,19	1,59	1,55	1,41
pvdS1	245,11	256,3	389,16	344,91	1,05	1,59	1,41
Cp1002_0672	111,77	166,94	144,63	156,48	1,49	1,29	1,40
pspA1	1030,56	1127,19	1234,47	1441,43	1,09	1,20	1,40
Cp1002_1427	1052,82	1049,73	1207,73	1472,18	1,00	1,15	1,40
rpmJ	539,97	2473,06	1362,12	754,92	4,58	2,52	1,40
trmA	86,61	191,27	171,85	120,65	2,21	1,98	1,39
Cp1002_1094	42,14	67,5	87,69	58,64	1,60	2,08	1,39
pip	157,98	205,11	233,68	219,65	1,30	1,48	1,39
bioD	97,02	144,3	117,81	134,52	1,49	1,21	1,39
Cp1002_0519	4582,65	926,82	2932,54	6331,66	0,20	0,64	1,38
thrS	681,42	1045,46	688,66	939,39	1,53	1,01	1,38
Cp1002_1032	122,63	153,87	183,77	169,04	1,25	1,50	1,38
purE	157,56	227,1	199,96	216,59	1,44	1,27	1,37
tgt	177,96	210,47	196,09	244,28	1,18	1,10	1,37
ftsI	131,54	415,9	254,37	180,22	3,16	1,93	1,37
hisF1	77,38	186,29	133,25	105,93	2,41	1,72	1,37
miaB	159,89	418,83	227,45	218,19	2,62	1,42	1,36
inv	196,39	310,8	242,15	267,81	1,58	1,23	1,36
pspA	1784,31	1830,96	1983,33	2421,97	1,03	1,11	1,36
rsmE	144,75	453,07	241,95	196,06	3,13	1,67	1,35
Cp1002_1861	159,08	278,06	192,92	215,37	1,75	1,21	1,35
Cp1002_0089	252,07	392,65	425,7	340,92	1,56	1,69	1,35
gpmB	183,66	384,21	285,55	247,43	2,09	1,55	1,35
Cp1002_1613	84,92	358,98	113,45	114,35	4,23	1,34	1,35
Cp1002_0131	122,58	211,72	203,35	163,27	1,73	1,66	1,33
Cp1002_0446	514,68	552,56	620,41	684,12	1,07	1,21	1,33
secF	169,34	313,01	306,34	224,7	1,85	1,81	1,33
Cp1002_0007	123,9	161,04	295,15	163,82	1,30	2,38	1,32
coaBC	56,44	134,6	86,77	74,48	2,38	1,54	1,32
Cp1002_2033	157,27	177,61	188,71	206,94	1,13	1,20	1,32
Cp1002_0275	74,46	120,53	97,83	97,81	1,62	1,31	1,31
xerD	74,1	184,43	133,91	97,21	2,49	1,81	1,31

continua

Cp1002_1660	264,6	464,02	433,15	346,83	1,75	1,64	1,31
coaD	222,66	324,97	341,44	291,45	1,46	1,53	1,31
rpoZ	704,4	1675,01	1607,53	915,23	2,38	2,28	1,30
proS	387,4	583,01	449,65	502,39	1,50	1,16	1,30
rplC	3211,71	3639,54	2843,65	4164,34	1,13	0,89	1,30
ccsB	112,4	205,25	196,93	145,71	1,83	1,75	1,30
recN	65,37	128,22	93,82	84,57	1,96	1,44	1,29
ftsQ	251,07	305,56	298,82	324,57	1,22	1,19	1,29
uvrA	141,18	255,51	199,95	182,4	1,81	1,42	1,29
yceG	243,53	457,87	348,75	314,33	1,88	1,43	1,29
Cp1002_0857	371,7	361,79	374,55	479,65	0,97	1,01	1,29
cspA	7330,71	10594,49	8384,92	9428,89	1,45	1,14	1,29
Cp1002_1207	83,73	161,67	128,95	107,65	1,93	1,54	1,29
Cp1002_0702	548,89	750,22	823,71	705,55	1,37	1,50	1,29
Cp1002_0284	302,44	426,75	373	388,51	1,41	1,23	1,28
Cp1002_1013	118,62	208,96	244,25	151,91	1,76	2,06	1,28
dapA	557,62	995,82	876,96	713,84	1,79	1,57	1,28
rpsO	3986,21	4128,61	5576,15	5090,61	1,04	1,40	1,28
uvrB	251,72	389,78	334,75	321,08	1,55	1,33	1,28
Cp1002_0783	158,78	248,13	279,87	202,44	1,56	1,76	1,27
ubiE	521,81	688,74	559,27	663,42	1,32	1,07	1,27
csdA	402,6	577,66	461,48	511,81	1,43	1,15	1,27
rpsD	905,54	1010,9	972,87	1146,52	1,12	1,07	1,27
guaA	965,18	1147,57	1091,03	1218,68	1,19	1,13	1,26
Cp1002_0770	180,11	284,19	201,94	226,95	1,58	1,12	1,26
rplU	5351,48	19454,34	8818,61	6737,57	3,64	1,65	1,26
Cp1002_1208	170,12	334,98	290,27	213,69	1,97	1,71	1,26
Cp1002_0088	383,51	1063,09	553,98	480,49	2,77	1,44	1,25
nrdI	99,43	114,85	147,22	124,49	1,16	1,48	1,25
Cp1002_1721	1892,91	1899,29	3169,21	2347,61	1,00	1,67	1,24
rpiB	1070,42	1271,58	1064,03	1325,71	1,19	0,99	1,24
fmt	141,59	182,17	172,97	174,79	1,29	1,22	1,23
Cp1002_1168	240,85	445,22	337,16	296,41	1,85	1,40	1,23
pimA	115,7	169,95	162,63	142,2	1,47	1,41	1,23
pimT	122,28	191,54	149,27	150,28	1,57	1,22	1,23
pyrD	174,13	298,23	186,74	213,42	1,71	1,07	1,23
Cp1002_1068	668,82	802,98	800,43	819,58	1,20	1,20	1,23
cyoE	287,66	489,3	454,17	351,8	1,70	1,58	1,22
ahpD	523,66	573,13	1115,29	638,57	1,09	2,13	1,22
aspA	253,83	120,13	116,97	308,96	0,47	0,46	1,22
Cp1002_0671	360,13	396,64	434,77	438,28	1,10	1,21	1,22
ectP	107,51	292,66	175,89	130,77	2,72	1,64	1,22
folK	161,17	202,08	212,11	195,34	1,25	1,32	1,21
SbcD	105,09	128,99	140,98	127,36	1,23	1,34	1,21
gyrA	480,73	478,76	476,3	581,95	1,00	0,99	1,21
groEL	5094	4134,5	5760,56	6115,36	0,81	1,13	1,20
zwf	183,17	262,77	266,95	219,86	1,43	1,46	1,20
ppnk	111,15	270,34	151,13	133,28	2,43	1,36	1,20
Cp1002_1084	90,07	239,97	149,23	107,84	2,66	1,66	1,20
mihF	1186,25	2321,72	2146,59	1417,95	1,96	1,81	1,20
folE	188,22	280,09	265,41	224,94	1,49	1,41	1,20
purB	195,08	208,63	206,31	232,44	1,07	1,06	1,19
ssb1	3320,84	4109,27	4485,72	3955,13	1,24	1,35	1,19
Cp1002_0054	122,88	254,2	142,05	145,94	2,07	1,16	1,19
Cp1002_0552	336,69	439,72	432,73	398,29	1,31	1,29	1,18

Cp1002_1954	535,12	536,04	589,92	630,85	1,00	1,10	1,18
metK	250,07	382,24	317,83	293,63	1,53	1,27	1,17
pafA	83,27	133,5	126,36	97,38	1,60	1,52	1,17
Cp1002_1175	300,03	795,44	502,34	349,86	2,65	1,67	1,17
Cp1002_1953	564,37	626,36	694,99	658,07	1,11	1,23	1,17
whiP	183,48	217,66	251,62	213,94	1,19	1,37	1,17
hemE	208,78	328,18	344,1	243,32	1,57	1,65	1,17
Cp1002_1768	542,09	529,91	614,77	631,35	0,98	1,13	1,16
rpmB2	1440,25	1965,9	1520,73	1669,67	1,36	1,06	1,16
frr	858,55	1530,09	1732,38	989,73	1,78	2,02	1,15
ppc	417,37	430,8	429,97	480,32	1,03	1,03	1,15
Cp1002_0804	666,53	604,22	674,25	765,65	0,91	1,01	1,15
efp	512,55	588,52	652,3	588,67	1,15	1,27	1,15
cysE	91,21	133,5	115,76	104,73	1,46	1,27	1,15
argS	424,19	610,37	468,12	486,62	1,44	1,10	1,15
rpmG	3243,3	4342,47	4647,79	3703,71	1,34	1,43	1,14
Cp1002_0539	210,79	301,82	271,39	240,54	1,43	1,29	1,14
Cp1002_0315	227,69	83,78	235,37	259,74	0,37	1,03	1,14
Cp1002_0267	1031,26	1934,11	2457,43	1176,26	1,88	2,38	1,14
rpsH	2827,06	3689,32	5145,59	3220,17	1,31	1,82	1,14
ccdA	120,75	185,26	162,66	137,14	1,53	1,35	1,14
Cp1002_1015	1189,92	1337,47	1268,68	1344,82	1,12	1,07	1,13
Cp1002_0532	2809,1	4131,04	4467,74	3174,33	1,47	1,59	1,13
rpmC	2467,47	2059,7	2063,69	2785,73	0,83	0,84	1,13
asd	734,31	659,21	730,27	828,61	0,90	0,99	1,13
Cp1002_1515	4849,49	7715,36	5164,47	5469,7	1,59	1,06	1,13
sbm	408,55	209,1	222,16	457,95	0,51	0,54	1,12
folC	313,13	469,85	438,71	349,88	1,50	1,40	1,12
ftsX	512,38	671,87	663,62	572,44	1,31	1,30	1,12
etfB	300,62	367,25	341,82	334,78	1,22	1,14	1,11
gmk	1808,99	2880,48	2731,77	2000,08	1,59	1,51	1,11
glT	244,78	429,75	308,34	270,33	1,76	1,26	1,10
gatA	209,69	269,43	248,52	231,21	1,28	1,19	1,10
alaS	356,21	555,81	453,18	392,01	1,56	1,27	1,10
Cp1002_0096	158,41	181,83	174,95	173,99	1,15	1,10	1,10
dnaA	134	228,07	187,78	146,58	1,70	1,40	1,09
Cp1002_0071	419,51	201,43	454,38	458,87	0,48	1,08	1,09
srtA	263,5	259,76	391,71	287,8	0,99	1,49	1,09
pks	286,05	392,57	393,23	310,9	1,37	1,37	1,09
hspR	243,76	263,14	1094,19	264,55	1,08	4,49	1,09
ppx1	193,09	252,64	206,94	209,41	1,31	1,07	1,08
pnp	1062,57	1802,65	1394,12	1149,53	1,70	1,31	1,08
rph	542,94	599,69	633,78	587,02	1,10	1,17	1,08
ppx2	303,89	858,18	454,15	327,86	2,82	1,49	1,08
ndh	609,7	547,3	707,92	655,05	0,90	1,16	1,07
Cp1002_2094	151,82	179,22	195,71	163,02	1,18	1,29	1,07
lpd	1783,78	2979,28	2717	1913,22	1,67	1,52	1,07
aroC	205,58	267,04	220,54	220,22	1,30	1,07	1,07
Cp1002_1416	241,07	319,87	315,17	257,52	1,33	1,31	1,07
rpsB	5934,68	6772,75	7632,89	6317,11	1,14	1,29	1,06
pqqE	262,31	344,17	315,05	278,42	1,31	1,20	1,06
atpH	1843,14	3034,23	2478,06	1950,56	1,65	1,34	1,06
vap	323,48	1260,54	358,79	342,3	3,90	1,11	1,06
sdaA	1475,45	607,06	540,2	1557,61	0,41	0,37	1,06
rplQ	5120,27	5477,33	5901,93	5402,96	1,07	1,15	1,06

rnc	1006,78	911,46	942,68	1061,71	0,91	0,94	1,05
Cp1002_2074	159,24	275,98	216,96	167,66	1,73	1,36	1,05
gyrB	898,29	1187,78	937,9	944,61	1,32	1,04	1,05
Cp1002_0106	143,46	181,07	165,33	150,77	1,26	1,15	1,05
nrde	183,23	192,85	233,49	192,45	1,05	1,27	1,05
Cp1002_1523	301,22	587,09	444,23	316,06	1,95	1,47	1,05
rpsQ	4197,15	4415,02	4560,52	4402,02	1,05	1,09	1,05
Cp1002_0058	1719,87	1765,98	1681,39	1799,51	1,03	0,98	1,05
mmdA	1341,59	466,53	704,05	1403,55	0,35	0,52	1,05
ddl	363,93	353,47	398,59	380,17	0,97	1,10	1,04
rplR	2549,87	2937,28	3272,4	2646,33	1,15	1,28	1,04
murC	274,44	357,29	343,55	283,26	1,30	1,25	1,03
Cp1002_0247	198,3	248,81	328,43	204,07	1,25	1,66	1,03
tuf	7045,66	5607,45	7045,2	7237,92	0,80	1,00	1,03
pheT	324,2	462,44	418,34	332,5	1,43	1,29	1,03
Cp1002_1151	180,25	286,61	228,62	184,78	1,59	1,27	1,03
Cp1002_1261	177,36	205,65	247,49	181,58	1,16	1,40	1,02
sdhA	883,84	935,4	511,77	903,5	1,06	0,58	1,02
ispG	449,96	476,95	539,51	459,52	1,06	1,20	1,02
pyrE	235,31	253,16	298,41	240,26	1,08	1,27	1,02
apt	369,22	716,09	462,68	375,94	1,94	1,25	1,02
whiA	250,2	277,34	328	254,66	1,11	1,31	1,02
tsf	2944,23	3430,16	4391,06	2988,09	1,17	1,49	1,01
atpC	11411,9	8966,96	5237,38	11577,8	0,79	0,46	1,01
pheS	191,49	387,95	243,72	193,2	2,03	1,27	1,01
hisG	209,14	242,3	241,59	210,58	1,16	1,16	1,01
murF	331,69	472,27	452,13	332,06	1,42	1,36	1,00
purM	533,12	671,58	498,76	531,72	1,26	0,94	1,00
rplS	1910,17	2232,52	2375,35	1892,64	1,17	1,24	0,99
atpB	905,6	1592,54	1451,23	895,47	1,76	1,60	0,99
Cp1002_0422	383,3	569,83	603,81	378,26	1,49	1,58	0,99
Cp1002_0661	2046,84	3551,14	2478,75	2018,4	1,73	1,21	0,99
ftsW	310,66	402,86	378,9	306,08	1,30	1,22	0,99
secG	551,29	591,16	808,95	542,13	1,07	1,47	0,98
rpoC	2154,12	1985,39	2111,19	2113,78	0,92	0,98	0,98
valS	357,6	453,44	421	349,33	1,27	1,18	0,98
ribA	329,21	528,05	446,67	321,28	1,60	1,36	0,98
sigA	259,94	582,09	410,5	252,56	2,24	1,58	0,97
Cp1002_1318	244,32	346,69	272,38	236,65	1,42	1,11	0,97
atpA	2410,53	2994,06	2611,51	2334,72	1,24	1,08	0,97
manB	685,23	716,47	672,95	661,51	1,05	0,98	0,97
Cp1002_1080	525,3	266,01	505,04	506,7	0,51	0,96	0,96
Cp1002_0963	482,2	672,2	518,41	463,46	1,39	1,08	0,96
Cp1002_0956	1020,38	1934,98	1260,02	975,96	1,90	1,23	0,96
pccB	536,34	693,73	529,43	511,71	1,29	0,99	0,95
Cp1002_0735	365,84	482,72	443,35	348,35	1,32	1,21	0,95
rpmH	1085,02	2781,72	1895,14	1023,47	2,56	1,75	0,94
Cp1002_0962	333,3	694,99	552,11	313,78	2,09	1,66	0,94
Cp1002_1185	351,94	512,89	379,95	329	1,46	1,08	0,93
rpmF	3238,24	3950,33	4370,99	2987,8	1,22	1,35	0,92
rpoB	849,18	952,12	1074,85	782,1	1,12	1,27	0,92
atpF	2570,3	3518,24	3157,94	2342,22	1,37	1,23	0,91
Cp1002_0663	822,04	464,66	753,26	747,7	0,57	0,92	0,91
aroH	443,67	460,33	433,15	402,27	1,04	0,98	0,91
ppa	611,74	589,41	672,3	546,64	0,96	1,10	0,89

Cp1002_1069	758,86	767,14	822,68	674,5	1,01	1,08	0,89
nusB	789,25	806,83	826,42	694,83	1,02	1,05	0,88
pflAC	2363,39	1186,57	1144,63	2071,58	0,50	0,48	0,88
malE	2521,77	1649,29	1711,53	2190,12	0,65	0,68	0,87
Cp1002_1193	3009,55	1952,93	2967,28	2580,22	0,65	0,99	0,86
clpC	1941,04	1070,31	1957,91	1663,83	0,55	1,01	0,86
rpsE	2695,08	3350,28	2957,67	2298,03	1,24	1,10	0,85
rplO	3306,59	3934,74	4341,88	2812,79	1,19	1,31	0,85
rplF	4338,98	5855,41	5292,09	3682,46	1,35	1,22	0,85
atpG	4609,92	4605,16	3845,9	3863,71	1,00	0,83	0,84
Cp1002_0785	1085,66	642,59	1022,65	902,87	0,59	0,94	0,83
rpsC	3288,26	1814,51	2033,97	2706,49	0,55	0,62	0,82
groEL1	7421,01	6047,05	9662,71	5923,74	0,81	1,30	0,80
rplL	7513,85	7892,14	9397,76	5901,82	1,05	1,25	0,79
Cp1002_0504	4504	2559,18	3130,11	2870,21	0,57	0,69	0,64
ctaE	1109,06	1662,63	1349,24	661,34	1,50	1,22	0,60
rpsL	986,36	1273,95	1849,49	579,03	1,29	1,88	0,59
tcsR	1157,81	282,53	483,84	669,96	0,24	0,42	0,58
Cp1002_0716	788,28	1049,75	851,93	431,31	1,33	1,08	0,55
Cp1002_0585	2640,86	999,48	3588,95	1442,32	0,38	1,36	0,55
Cp1002_1689	387,68	585,14	570,76	206,83	1,51	1,47	0,53
glyA	852,89	421,68	538,7	452,39	0,49	0,63	0,53
mcmA	1267,26	669,68	814,68	658,43	0,53	0,64	0,52
Cp1002_0998	301,25	442,36	362,08	155,13	1,47	1,20	0,51
Cp1002_1165	394,77	384,92	457,9	195,74	0,98	1,16	0,50
yqgF	452,75	1057,44	463,31	224,48	2,34	1,02	0,50
nrdF	463,32	820,74	579,24	226,99	1,77	1,25	0,49
Cp1002_2078	364,13	151,03	186,85	177,58	0,41	0,51	0,49
Cp1002_1442	1583,77	1557,41	1379,98	742,91	0,98	0,87	0,47
sepF	1425,12	1445,97	1624,87	668,22	1,01	1,14	0,47
Cp1002_0643	195,19	311,46	219,33	91,49	1,60	1,12	0,47
tyrS	662,42	933,05	656,65	306,24	1,41	0,99	0,46
gatC	192,58	278,25	225,2	88,14	1,44	1,17	0,46
ppm1	945,54	864,3	928,54	429,58	0,91	0,98	0,45
Cp1002_0851	1022,38	319,58	544,38	461,05	0,31	0,53	0,45
carD	280,62	431,37	467,91	126,15	1,54	1,67	0,45
Cp1002_2011	6575	4298,77	5444,98	2881,3	0,65	0,83	0,44
Cp1002_1000	216,85	218,48	241,4	93,94	1,01	1,11	0,43
mdh	2924	1453,1	1834,62	1261,86	0,50	0,63	0,43
plsC	1120,16	1516,21	1401,05	481,35	1,35	1,25	0,43
pycB	2918,89	864,8	1816,18	1250,55	0,30	0,62	0,43
groES	10603,88	6335,14	9023,8	4538,76	0,60	0,85	0,43
Cp1002_0849	1149,29	569,88	1249,18	489,68	0,50	1,09	0,43
Cp1002_0573	1902,87	453,09	824,5	768,55	0,24	0,43	0,40
Cp1002_2049	390,15	191,98	412,61	152,96	0,49	1,06	0,39
argJ	246,31	98,54	113,77	96,21	0,40	0,46	0,39
Cp1002_1495	519,54	1566,09	836,01	198,44	3,01	1,61	0,38
rimM	227,91	284,51	308,43	85,69	1,25	1,35	0,38
ptsI	2304,11	1224,48	1489,02	836,71	0,53	0,65	0,36
Cp1002_1206	642,43	281,91	339,71	231,66	0,44	0,53	0,36
Cp1002_0954	650,54	981,69	626,43	230,88	1,51	0,96	0,35
Cp1002_0613	144,85	60,49	63,97	50,51	0,42	0,44	0,35
argH	653,73	366,74	378,32	225,31	0,56	0,58	0,34
oppB	151,03	28,23	45,3	50,3	0,19	0,30	0,33
cinA	150,87	321,63	247,7	50,22	2,13	1,64	0,33

pyrH	1007,73	1162,53	1206,54	328,19	1,15	1,20	0,33
rpfB	219,27	388,12	266,88	68,42	1,77	1,22	0,31
argG	971,12	343,32	589,34	295,67	0,35	0,61	0,30
dcuB	551,86	142,47	297,84	160,09	0,26	0,54	0,29
nanK	116,02	23,64	43,42	33,63	0,20	0,37	0,29
nusG	604,48	634,28	581,58	162,93	1,05	0,96	0,27
Cp1002_0432	131,58	170,83	152,8	34,97	1,30	1,16	0,27
orn	492,43	215,84	287,87	129,24	0,44	0,58	0,26
oppA	568,55	117,32	200,43	143,03	0,21	0,35	0,25
Cp1002_0174	190,54	43,1	76,77	44,99	0,23	0,40	0,24
glxR	1604,55	435,09	970,51	339,38	0,27	0,60	0,21
sdaC	1674,41	623,44	857,74	328,86	0,37	0,51	0,20
Cp1002_2063	351,02	77,45	105,96	66,01	0,22	0,30	0,19
Cp1002_1302	410,22	421,59	455,46	67,16	1,03	1,11	0,16
Cp1002_1908	247,26	19,25	107,74	21,05	0,08	0,44	0,09
Cp1002_0008	18,77	0	0	0	0	0	0

Genes comuns, significativos, entre as três condições testadas independente do valor de *fold-change*. *Fold-change* acima de 1, indica que gene na condição de estresse foi induzido em relação a condição controle. Abaixo de 1, o gene foi reprimido.

Anexo 9- Relatório do Processo osmótico induzido

Anexo 10- Relatório do Processo osmótico reprimido

Anexo 11- Relatório do Processo meio ácido induzido

Anexo 12- Relatório do Processo meio ácido reprimido

Anexo 13 - Relatório do Processo térmico induzido

Anexo 14 - Relatório do Processo térmico reprimido

Anexo 15 – Lista com todos os genes dos trabalhos do grupo LGCM envolvendo *C.pseudotuberculosis* 1002 e valores de *fold-change* para cada estresse

Anexo 16: Atividades Desenvolvidas (2007-2011)

Atuação profissional

- Disciplinas Ministradas: *Montagem e Anotação de Genoma*. Laboratório de Bioinformática do Instituto de Ciências Biológicas da UFPA- Belem-PA
Período: Junho/Julho de 2009
- Disciplinas Ministradas: *Montagem e Anotação de Genoma*. Instituto de Ciências Biológicas da UFOPA- Santarém-PA
Período: Dezembro/2009

Cursos complementares

- Curso de Estatística R -UFMG
Período: Setembro a Dezembro de 2008

Realização de experimentos científicos em colaboração com outras instituições de ensino e pesquisa

- Universidade Federal do Pará - Data: de junho de 2009 a julho de 2009.
- Universidade Federal do Pará - Data: de 01 novembro de 2009 a 18 de dezembro de 2009.
- Universidade Federal do Oeste do Pará (Santarém) - Data: de 07 de dezembro a 11 de dezembro de 2009

Participação em eventos

- 26 ° Congresso Brasileiro de Microbiologia (2011)
- 55° Congresso Brasileiro de Genética. Guarujá (2010)
- 27ª Reuniao de Genética de Microrganismos. Guaruja (2010)
- 25° Congresso Brasileiro de Microbiologia (2009).
- I Simpósio de Genética e Biotecnologia da UFMG - I Encontro de Alunos e Ex-alunos do PG Genética - Professora Cleusa Graça da Fonseca. Belo Horizonte (2008)
- 54° Congresso Brasileiro de Genética. Salvador (2008).

- 26ª Reunião de Genética de Microrganismos – Salvador (2008).
- IV Fórum de Microbiologia Professor Eduardo Osório Cisalpino-UFMG (2008)

Trabalhos publicados em anais de eventos (resumo)

TRABALHO SELECIONADO PARA APRESENTAÇÃO ORAL: PINTO, AC, SÁ, PHCG, RAMOS, RTJ, BARBOSA, S, SILVA, WM, ROCHA, FS, SCHNEIDER MP, MIYOSHI A, SILVA, A, AZEVEDO, V. Análise de expressão diferencial, em meio simulando acidez de macrófago, de *Corynebacterium pseudotuberculosis* utilizando RNAseq. In: 26º Congresso Brasileiro de Microbiologia, 2011, Foz do Iguaçu. Anais do 26º Congresso Brasileiro de Microbiologia., 2011.

SÁ, PHCG, PINTO, AC, RAMOS, RTJ, SANTOS, AV, AZEVEDO, V, SCHNEIDER, MPC, SILVA, A DBTransProt – Um Banco de Dados para Integração de Dados de Transcriptoma e Proteoma. In: 26º Congresso Brasileiro de Microbiologia, 2011, Foz do Iguaçu. Anais do 26º Congresso Brasileiro de Microbiologia., 2011

PINTO, A.C; D'AFONSECA, V.; SANTOS, A.R.; ALMEIDA, S.S.; SOARES, S.C.; FARIA, C.J; MAGALHÃES, A; RUIZ, J.C.; MIYOSHI, A, AZEVEDO, V. Estudo da Arquitetura Genômica de duas linhagens do patógeno *Corynebacterium Pseudotuberculosis* e seu estilo de vida. In: 55º Congresso Brasileiro de Genética. Guarujá (2010)

PINTO, A.C; D'AFONSECA, V.; SANTOS, A.R.; ALMEIDA, S.S.; SOARES, S.C.; FARIA, C.J; MAGALHÃES, A; RUIZ, J.C.; MIYOSHI, A, AZEVEDO, V. Estudo da Arquitetura Genômica de duas linhagens do patógeno *Corynebacterium Pseudotuberculosis* e seu estilo de vida. In: 27ª Reunião de Genética de Microrganismos. Guarujá (2010)

PINTO, A. C., RESENDE, D. M., SOARES, S., DAFONSECA, V., ALMEIDA, S., MOORE, R., SILVA, A., RUIZ, J. C., MIYOSHI, A., AZEVEDO, V.
ANALISE COMPARATIVA DOS PSEUDOGENES ENTRE DUAS LINHAGENS DE *Corynebacterium pseudotuberculosis* (C231 e 1002) E OUTRA ESPÉCIE *Corynebacterium diphtheriae* (NCTC 13129) In: 25º Congresso Brasileiro de Microbiologia, 2009, Porto de Galinhas. Anais do 25º Congresso Brasileiro de Microbiologia. , 2009.

DAFONSECA, V., ALMEIDA, S., SOARES, S. C., PINTO, A. C., Santos AR, RESENDE, D. M., OLIVEIRA, G., RAMOS, R. T. J., CERDERIA, L., CUADORS, S., SILVA, A., MIYOSHI, A., RUIZ, J. C., AZEVEDO, V. Comparative genomics between *Corynebacterium pseudotuberculosis* strains In: Simposio Interdisciplinar Física+Bioinformática, 2009, Ouro Preto, MG. Anais do Simposio Interdisciplinar Física+Bioinformática. , 2009.

SOARES, S., DAFONSECA, V., PINTO, A. C., ALMEIDA, S., CUADORS, S., SILVA, A., OLIVEIRA, G. C., RUIZ, J. C., AZEVEDO, V., MIYOSHI, A. Identificação, caracterização e comparação in silico das ilhas de patogenicidade (PAIS) de *Corynebacterium pseudotuberculosis* In: 25º congresso brasileiro de microbiologia, 2009, Porto de Galinhas. Anais do 25º congresso brasileiro de microbiologia. , 2009.

SOARES, S. C., DAFONSECA, V., ALMEIDA, S., PINTO, A. C., Santos AR, RAMOS, R. T. J., RESENDE, D. M., CUADORS, S., SILVA, A., MIYOSHI, A., AZEVEDO, V., RUIZ, J. C. Pathogenicity islands in bacteria: A next generation sequence approach overview of *Corynebacterium* genomes In: Simposio Interdisciplinar Fisica+Bioinformatica, 2009, ouro preto. Anais do Simpósio Interdisciplinar Física+Bioinformatica. , 2009.

RAMOS, R. T. J., DAFONSECA, V., ALMEIDA, S., CERDERIA, L., SOARES, S., PINTO, A. C., RESENDE, D. M., CUADORS, S., LAMEIRA, J., AZEVEDO, V., SILVA, A., RUIZ, J. C. Sequenciamento dos genomas de linhagens de *Corynebacterium pseudotuberculosis* utilizando a plataforma solid in: 25º congresso brasileiro de microbiologia, 2009, Porto de Galinhas. Anais do 25º congresso brasileiro de microbiologia. , 2009.

SOARES, S., PINTO, A. C., DORELLA, FA, Cerqueira, PG, PACHECO, LGC, SEYFFERT, N., DAFONSECA, V., CASTRO, T. L. P., HIRATA JUNIOR, R., AZEVEDO, V., MIYOSHI, A. Análise das ilhas de patogenicidade de *Corynebacterium diphtheriae* por PCR2 (plasticity of chromosome revealed by long range-polymerase chain reaction) In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. 2008.

SEYFFERT, N., PACHECO, LGC, CASTRO, T. L. P., DORELLA, FA, Cerqueira, PG, PINTO, A. C., SOARES, S., DAFONSECA, V., Oliveira, SC, PIMENTA, A., MEYER, R., MIYOSHI, A., AZEVEDO, V. Análise imunoproteômica do secretoma de *Corynebacterium pseudotuberculosis* para fins diagnósticos. In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. 2008.

DORELLA, FA, Cerqueira, PG, PACHECO, LGC, SEYFFERT, N., PINTO, A. C., DAFONSECA, V., CASTRO, T. L. P., SOARES, S., Oliveira, SC, MEYER, R., MIYOSHI, A., AZEVEDO, V. Analysis of the vaccine potential of *Corynebacterium pseudotuberculosis* mutants obtained through random mutagenesis In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. 2008.

DORELLA, FA, Cerqueira, PG, PACHECO, LGC, SEYFFERT, N., PINTO, A. C., DAFONSECA, V., CASTRO, T. L. P., SOARES, S., Oliveira, SC, MEYER, R., MIYOSHI, A., AZEVEDO, V. Analysis of the vaccine potential of *Corynebacterium pseudotuberculosis* mutants obtained through random mutagenesis. In: 54º Congresso Brasileiro de Genética, 2008, salvador. Anais do 54º Congresso Brasileiro de Genética, 2008. 2008.

DAFONSECA, V., Morais, PMRO, DORELLA, FA, PACHECO, LGC, ALMEIDA, S., PINTO, A. C., Santos AR, Cerqueira, PG, SEYFFERT, N., CASTRO, T. L. P., SOARES, S., Oliveira, SC, ORTEGA, J. M., MEYER, R., MIYOSHI, A., AZEVEDO, V. Caracterização do conteúdo gênico e organização genômica de *Corynebacterium pseudotuberculosis* In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. 2008.

Morais, PMRO, DORELLA, FA, PACHECO, LGC, ALMEIDA, S., Santos AR, PINTO, A. C., Cerqueira, PG, SEYFFERT, N., CASTRO, T. L. P., AZEVEDO, V. Construção de uma biblioteca genômica de *Corynebacterium pseudotuberculosis* e sua caracterização através de Genome Sequence Survey (GSS). In: 26 Reunião de Genética de Microrganismos, 2008, Salvador-Ba. 26 Reunião de Genética de Microrganismos. 2008.

Cerqueira, PG, PINTO, A. C., DORELLA, FA, PACHECO, LGC, SOARES, S., CASTRO, T. L. P., SEYFFERT, N., MEYER, R., MIYOSHI, A., AZEVEDO, V. Construção de uma biblioteca genômica de *Corynebacterium pseudotuberculosis* para a identificação de promotores através da expressão de gfp In: 54º Congresso Brasileiro de Genética, 2008, Salvador. 54º Congresso Brasileiro de Genética. , 2008.

Cerqueira, PG, PINTO, A. C., DORELLA, FA, PACHECO, LGC, SOARES, S., CASTRO, T. L. P., DAFONSECA, V., SEYFFERT, N., MIYOSHI, A., AZEVEDO, V. Construção de uma biblioteca genômica de *Corynebacterium pseudotuberculosis* para a identificação de promotores através da expressão do gfp In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. , 2008.

SILVA, W., SEYFFERT, N., PACHECO, LGC, CASTRO, T. L. P., SOARES, S., DORELLA, FA, Cerqueira, PG, PINTO, A. C., DAFONSECA, V., MIYOSHI, A., AZEVEDO, V. Identificação de epítomos da superfície de fagos para o imunodiagnóstico da LC em caprinos In: 54º Congresso Brasileiro de Genética, 2008, Salvador. 54º Congresso Brasileiro de Genética.2008.

SEYFFERT, N., PACHECO, LGC, CASTRO, T. L. P., DORELLA, FA, Cerqueira, PG, PINTO, A. C., SOARES, S., Morais, PMRO, DAFONSECA, V., Oliveira, SC, PIMENTA, A., MEYER, R., MIYOSHI, A., AZEVEDO, V. Integração da genômica e proteômica de *Corynebacterium pseudotuberculosis* para o imunodiagnóstico da LC. In: 54º Congresso Brasileiro de Genética, 2008, Salvador. 54º Congresso Brasileiro de Genética. , 2008.

PACHECO, LGC, CASTRO, T. L. P., SEYFFERT, N., DORELLA, FA, PINTO, A. C., SOARES, S., BUCKER, D., Morais, PMRO, Cerqueira, PG, SILVA, W., MEYER, R., MIYOSHI, A., Oliveira, SC, AZEVEDO, V. Papel do fator sigma E de *Corynebacterium pseudotuberculosis* na resistência ao estresse e na regulação de fatores associados à virulência In: 54º Congresso Brasileiro de Genética, 2008, Salvador. 54º Congresso Brasileiro de Genética. , 2008.

PACHECO, LGC, CASTRO, T. L. P., SEYFFERT, N., DORELLA, FA, PINTO, A. C., DAFONSECA, V., SOARES, S., BUCKER, D., Morais, PMRO, COELHO, A. M., SILVA, W., PIMENTA, A., MEYER, R., MIYOSHI, A., Oliveira, SC, AZEVEDO, V. Papel do fator sigma E de *Corynebacterium pseudotuberculosis* na resistência ao estresse e na regulação de fatores associados à virulência. In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. , 2008.

SOARES, S., PINTO, A. C., DORELLA, FA, Cerqueira, PG, PACHECO, LGC, SEYFFERT, N., CASTRO, T. L. P., HIRATA JUNIOR, R., AZEVEDO, V., MIYOSHI, A. Plasticidade genômica de *Corynebacterium diphtheriae*: implicações na virulência In: 54º Congresso Brasileiro de Genética, 2008, Salvador. 54º Congresso Brasileiro de Genética. , 2008.

SILVA, W., SEYFFERT, N., PACHECO, LGC, CASTRO, T. L. P., DORELLA, FA, Cerqueira, PG, PINTO, A. C., DAFONSECA, V., SOARES, S., MEYER, R., MIYOSHI, A., AZEVEDO, V. Uso de phage display na seleção de peptídeos para o diagnóstico da LC em pequenos ruminantes In: 26ª Reunião de Genética de Microrganismos, 2008, Salvador. 26ª Reunião de Genética de Microrganismos. , 2008.

Trabalhos publicados em anais de eventos (resumo expandido)

SIOMAR DE CASTRO SOARES, ARTUR LUIZ DA COSTA SILVA, ROMMEL THIAGO JUCÁ RAMOS, LOUISE TEIXEIRA CERDEIRA, AMJAD ALI, ANDERSON RODRIGUES DOS SANTOS, ANNE CYBELLE PINTO, ARYANE APARECIDA MAGALHÃES CASSIANO ROCHA, EUDES GUILHERME VIEIRA BARBOSA, LAIANE ALVES LEÃO, SINTIA SILVA DE ALMEIDA, VINÍCIUS AUGUSTO CARVALHO DE ABREU, ANDERSON MIYOSHI, VASCO AZEVEDO. Plasticidade Genômica e Evolução Bacteriana. Microbiologia in foco. Revista Microbiologia in foco, p.8 – 2011

RAMOS, R. T. J., CERDERIA, L., DAFONSECA, V., ALMEIDA, S., RESENDE, D. M., PINTO, A. C., SOARES, S. C., Santos AR, SILVA, R. R., CUADORS, S., SCHNEIDER, M. P. C., LAMEIRA, J., OLIVEIRA, G. C., SILVA, A., AZEVEDO, V., RUIZ, J. C. A *pipeline* for bacterial genome sequencing assembly using short reads: *Corynebacterium pseudotuberculosis* as a model In: Simpósio Interdisciplinar Física + Bioinformática, 2009, Ouro Preto. Anais do Simpósio Interdisciplinar Física + Bioinformática. , 2009.

DAFONSECA, V., ALMEIDA, S., SOARES, S., PINTO, A. C., RAMOS, R. T. J., SILVA, A., MOORE, R., MIYOSHI, A., RUIZ, J. C., AZEVEDO, V. Genômica comparativa entre diferentes linhagens de *Corynebacterium pseudotuberculosis* In: 25º Congresso Brasileiro de Microbiologia, 2009, Porto de Galinhas-PE.

RUIZ, J. C., SANTOS. A.R, PINTO, A. C., RESENDE, D. M., CERDERIA, L., RAMOS, R. T. J., CUADORS, S., ALMEIDA, S., SOARES, S. C., DAFONSECA, V., AZEVEDO, V., SILVA, A. Segunda Revolução Genômica: utilização de sequenciadores de nova geração. Microbiologia in foco. Revista Microbiologia in foco, p.15 - 18, 2009.