UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOINFORMÁTICA

Izabela Coimbra Ibraim

Análise do transcriptoma de *Corynebacterium pseudotuberculosis* em resposta à limitação por ferro

ORIENTADOR: Dr. Vasco A. C. Azevedo
COORIENTADORA: Dra. Anne Cybelle Pinto Gomide

BELO HORIZONTE
Novembro 2018

IZABELA COIMBRA IBRAIM

Análise do transcriptoma de *Corynebacterium pseudotuberculosis* em resposta à limitação de ferro

Tese apresentada ao programa de Pós-graduação em Bioinformática da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Bioinformática.

ORIENTADOR: Dr. Vasco A. C. Azevedo
COORIENTADORA: Dra. Anne Cybelle Pinto Gomide

BELO HORIZONTE
Novembro 2018

Para tudo há uma ocasião, e um tempo para cada
propósito debaixo do céu:
tempo de nascer e tempo de morrer,
tempo de plantar e tempo de arrancar o que se plantou,
tempo de matar e tempo de curar,
tempo de derrubar e tempo de construir,
tempo de chorar e tempo de rir,
tempo de prantear e tempo de dançar,
tempo de espalhar pedras e tempo de ajuntá-las,
tempo de abraçar e tempo de se conter,
tempo de procurar e tempo de desistir,
tempo de guardar e tempo de lançar fora,
tempo de rasgar e tempo de costurar,
tempo de calar e tempo de falar,
tempo de amar e tempo de odiar,
tempo de lutar e tempo de viver em paz.
Eclesiastes 3:1-8

AGRADECIMENTOS

Agradeço à Universidade Federal de Minas Gerais pela excelente formação acadêmica que me foi oferecida.

Ao Prof. Vasco Azevedo, por toda o apoio científico e por essa incrível oportunidade. O senhor foi além de suas obrigações como orientador. Sua ajuda e apoio foram para mim de valor inestimável e só tenho que agradecer a confiança depositava em mim.

À Dra. Anne Cybelle, por assumir a tarefa de coorientação deste trabalho, e por contribuir enormemente para a minha formação.

À Dra. Gloria Franco, por ser um grande exemplo de pessoa e professional. Por sua abertura, atenção, disponibilidade e comprometimento, o meu especial muito obrigada!

Ao Dr. Emanuel, pelo sequenciamento das amostras e por todo seu apoio científico.

Aos meus familiares que contribuíram decisivamente para o desenvolvimento deste trabalho. Dona Cida, Joseph, Jolie e Junior, obrigada por tudo!

Debs, só tenho que pedir desculpas por não ter lhe ajudado da mesma forma que me ajudou. You were and always will be my *bestest* friend in the whole world. Te amo! Rest in peace!

Ao meu marido Rafael Assis, palavras não podem descrever minha gratidão por tudo que fez por mim e pelo Henrique.

À Sheila Santana, não existem palavras suficientes para expressar o quanto sou grata pela sua ajuda. Obrigada por me escutar, aconselhar e me ajudar durante esses três anos. Você é simplesmente uma pessoa espetacular. Que Deus possa te retribuir em dobro tudo aquilo que fez por mim.

À secretaria do departamento de Bioinformática da UFMG, Marcia Natalia e Thiago, pela paciência, atenção e imensa ajuda na parte burocrática do processo.

Aos amigos e colegas do LGCM. Vocês estiverem comigo nos melhores e nos mais difíceis momentos da minha vida. Sem vocês eu não teria conseguido completar esse trabalho. Obrigada pela amizade e o companheirismo. Rock on my peeps!

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) que subsidiou o desenvolvimento desse projeto, meu muito obrigada.

Aos funcionários da biblioteca da UFMG.

Por fim, agradeço ao Programa de Pós-Graduação e a todos os professores do departamento pela imensa oportunidade.

Agradeço a Deus acima de qualquer coisa por essa vitória.

SUMÁRIO

I. LISTA DE FIGURAS 6
II. LISTA DE TABELAS 6
III. LISTA DE ABREVIATURAS 7
IV. RESUMO 8
V. ABSTRACT 10
VI. DELINEAMENTO 12
1. INTRODUÇÃO 13
1.1. *Corynebacterium pseudotuberculosis* 13
1.2. Homeostase de ferro em bactérias 17
1.2.1. Regulação da expressão gênica dependente de ferro em bactérias 18
1.2.2. Mecanismos de captação e armazenamento de ferro em bactérias 19
1.3. A pesquisa por determinantes de virulência em C. pseudotuberculosis 21
2. JUSTIFICATIVA 25
3. OBJETIVOS 27
3.1. Objetivo Geral 27
3.2. Objetivos Específicos 27
4. Metodologia 28
4.1. Linhagens bacterianas e condições de cultivo 28
4.2. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de *C. pseudotuberculosis* em meio (BHI) e quantificação da concentração de ferro no meio BHI . 28
4.3. Preparo do ligante de ferro 2,2-bipiridina (BIP) e determinação da concentração de uso 29
4.4. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de *C. pseudotuberculosis* em meio (BHI) suplementado com 250 µM do ligante 2,2-bipiridina 30
4.5. Desenho experimental para a análise de expressão diferencial em resposta à limitação de ferro 31
4.6. Estatísticas do sequenciamento e avaliação da qualidade das reads 33
5. RESULTADOS 34
5.1. Artigo: Transcriptome Profile of *Corynebacterium pseudotuberculosis* in Response to Iron Limitation (Submetido para publicação no periódico BMC Genomics – Online ISSN: 1471-2164). 34
5.1.1. Abstract 35
5.1.2. Background 38
5.1.3. Results and Discussion 41
5.1.4. Conclusion 57
5.1.5. Methods 58
5.1.6. Additional files 64

5.1.7. Abbreviations 66
5.1.8. Acknowledgement 67
5.1.9. Authors' contributions 67
5.1.10. Conflict of interest statement 67
5.1.11. Consent for publication 67
5.1.12. Ethics approval and consent to participate 67
5.1.13. References 68
5.1.14. Figure List: 78
5.2. Resultados Adicionais 84
5.2.1. Quantificação da concentração de ferro no meio BHI e avaliação do crescimento das linhagens T1 parental e Cp13 mutante de C. pseudotuberculosis em meio BHI 84
5.2.2. Preparo do ligante de ferro 2,2-bipiridina (BIP) e determinação da concentração de uso 85
5.2.3. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de C. pseudotuberculosis em meio (BHI) suplementado com 250 µM do ligante 2,2-bipiridina 87
5.2.4. Análise da qualidade do RNA total 89
5.2.5. Estatísticas do sequenciamento e avaliação da qualidade das reads 90
5.2.6. Avaliação da qualidade dos dados e identificação do efeito de batch 94
5. CONCLUSÕES 96
6. PERSPECTIVAS 97
7. REFERÊNCIAS BIBLIOGÁFICAS 99
8. ANEXOS 111
8.1. ANEXO I: Material suplementar do artigo apresentado na seção 4.8. 111
8.2. ANEXO II: Artigo publicado no periódico GENE (ISSN: 0378-1119) - Transcriptome analysis of *Corynebacerium pseudotuberculosis* biovar Equi in two condition of environmental stress (doi: 10.1016/j.gene.2018.08.028). 139
8.2.1. Introduction 140
8.2.2. Materials and Methods 143
8.2.3. Results and Discussion 147
8.2.4. Acknowledgements 158
8.2.5. Additional files 159
8.2.6. References 159
8.2.7. Tables 167
8.2.8. Figures 168
8.3. ANEXO III: Material suplementar do artigo apresentado no ANEXO II 176

I. LISTA DE FIGURAS

Figure 1. Estrutura do sistema TnFuZ 22
Figure 2. Representação esquemática do cluster ciu em *C. pseudotuberculosis*. 23
Figure 3. *C. pseudotuberculosis* growth assays. 78
Figure 4. Differential gene expression of *C. pseudotuberculosis* Cp13 mutant and T1 wild-type under iron limitation (LI/HI). 79
Figure 5. Protein interaction network analyses of up and down-regulated genes in *C. pseudotuberculosis* strains T1 and Cp13. 80
Figure 6. Expression pattern of commonly expressed genes identified in the T1 strain and the Cp13 mutant. 81
Figure 7. Genetic map of the target genes identified as having a putative iron-DtxR regulated binding site. 82
Figure 8. C. pseudotuberculosis transcriptional regulatory iron network 83
Figure 9. Curvas de crescimento e viabilidade bacteriana das linhagens selvagem (T1) e mutante (Cp13) em meio BHI. 85
Figure 10. Curvas de crescimento das linhagens selvagem (T1) e mutante (Cp13) em meio BHI suplementado com diferentes concentrações do ligante 2,2-bipiridina 86
Figure 11. Curvas de crescimento das linhagens selvagem (T1) e mutante (Cp13) em meio BHI -HI e meio BHI suplementado com 250 µM do ligante 2,2-bipiridina. 88
Figure 12. Análise qualitativa do perfil eletrofóretico dos RNAs totais extraídos de amostras da linhagem T1 selvagem e Cp13 mutante. 89
Figure 13. Valores de qualidade por base. 92
Figure 14. Análise por componentes principais (PCA) das amostras HI e LI da linhagem Cp13 mutante e T1 selvagem e distância Euclidiana das amostras da linhagem Cp13 mutante 95

II. LISTA DE TABELAS

Tabela 1: Overlapped differentially expressed genes identified in the *C. pseudotuberculosi*s Cp13 mutant and T1 strain. 76
Tabela 2: Estatística geral das bibliotecas sequenciadas de culturas HI (High Iron) e LI (Low Iron) das linhagens Cp13 mutante e T1 selvagem. 93

III. LISTA DE ABREVIATURAS

ATP – Adenosina trifosfato

BHI – Infusão cérebro coração

BIP ou DIP –2,2’-bipiridina ou 2,2-dipypirydyl

DO – Densidade ótica

FDR – False Discovery rate

FT – Fator de transcrição

GDEs – Genes diferencialmente expressos

GO – Gene ontology (ontologia gênica)

HI – Condição experimental de alta disponibilidade de ferro – *High Iron*

HMM – Modelo oculto de Markov

KEEG – Enciclopédia Kyoto de genes e genomas Kyoto

LC – Linfadenite caseosa

LI – Condição experimental de baixa disponibilidade de ferro – *Low Iron*

LU – Linfangite ulcerativa

mL – mililitro

pb – Pares de base

PCA –Análise de componentes principais

PCR – Reação em cadeia da polimerase

PFAM – Base de dados de família de proteínas

PLD – Fosfolipase D

rlog – Função para transformação logarítmica

RNS – Espécies reativas de nitrogênio

ROS – Espécies reativas de oxigênio

rpm – Rotações por minuto

RRT – Redes regulatórias transcricionais

TCA - Ciclo dos ácidos tricarboxílicos ou ciclo de Krebs

UFC – Unidades formadora de colônias

µl – microlitro

IV.RESUMO

O gene *ciuA* codifica uma proteína de superfície associada a um sistema de captação de ferro dependente de sideróforos, porém seu papel na aquisição de ferro nessa bactéria ainda é pouco conhecido. O ferro é crucial no crescimento e desenvolvimento de bactérias, porém no hospedeiro as frações livres disponíveis de ferro são inferiores às necessárias para o crescimento bacteriano. Em bactérias intracelulares como *C. pseudotuberculosis*, a homeostase de ferro no interior do hospedeiro é controlada transcricionalmente pela concentração de ferro no meio e envolve a regulação de genes voltados para a subversão dos sistemas de restrição nutricional imposto pelo hospedeiro. Trabalhos anteriores realizados por nosso grupo, levaram a construção de uma linhagem mutante (Cp13) para o gene *ciuA*, através de mutagênese aleatória utilizando o transposon TnFuZ na linhagem T1 selvagem de *Corynebacterium pseudotuberculosis*. A interrupção deste gene, contribuiu para a redução da virulência e viabilidade intracelular da linhagem, além de induzir níveis significativos de anticorpos e alta taxa de sobrevivência em camundongos imunizados. Diante do exposto, este trabalho teve como objetivo principal avaliar o crescimento e a resposta transcricional de *C. pseudotuberculosis* em resposta à restrição de ferro, bem como a caracterização funcional de genes associados a virulência, entre linhagem T1 e mutante Cp13. Assim, a análise de expressão diferencial de ambas as linhagens, foi realizada a partir de culturas incubadas em meio com baixa disponibilidade de ferro (*low iron* – LI), obtido pelo agente quelante 2,2-bipiridina (BIP), e meio com alta disponibilidade de ferro (*high iron* – HI). Neste contexto, os dados aqui apresentados demonstraram uma redução significativa na densidade celular e na taxa de crescimento da linhagem mutante em comparação a linhagem parental em culturas LI. No total, 120 e 86 genes diferencialmente expressos (GDEs) foram identificados nas cepas T1 e Cp13 em resposta à restrição de ferro, respectivamente. Dentre esses genes, 36 GDEs foram compartilhados entre as duas linhagens, sendo que 26 apresentaram um perfil de expressão semelhante entre as linhagens. Dois fatores transcricionais *ripA* e *hrrA*, regulados por DtxR, foram identificados entre os GDEs compartilhados, demonstrando uma complexa rede regulatória controlada pela disponibilidade de ferro. Sete dos 36 GDEs codificam proteínas de superfície e transportadores de hemina, sendo quatro genes localizados em ilhas genômicas associados a fatores de virulência. Nossos resultados demonstraram que a restrição intracelular da utilização de ferro é fundamental para a resposta de *C. pseudotuberculosis* à disponibilidade do metal. Alterações

significativas também foram observadas em relação aos GDEs identificados exclusivamente entre as linhagens, indicando divergência entre a resposta adaptativa da linhagem mutante em relação a linhagem parental. Em conjunto, os resultados do presente trabalho fornecem uma análise abrangente da rede regulatória e da resposta transcricional da linhagem T1 selvagem e Cp13 mutante de *C. pseudotuberculosis*, acrescentando informações importantes envolvendo o processo de adaptação desse importante patógeno à restrição de ferro.

Palavras-chave: *Corynebacterium pseudotuberculosis*, homeostase de ferro, fatores transcricionais regulados por ferro, expressão gênica diferencial

V. ABSTRACT

The *ciuA* gene codes for a surface protein which is associated to a siderophore-base iron acquisition system; however, its role in the acquisition of iron in *C. pseudotuberculosis* is still largely unknown. Iron is crucial for the growth and development of many bacterial pathogens; however, within the host, the concentration that is freely accessible is much lower than the concentration required for bacterial growth. In intracellular bacteria, like *C. pseudotuberculosis*, iron homeostasis is transcriptionally controlled by the availability of iron in the environment, which allows the bacteria to counteract the limitation of the metal imposed by the host. In a previous work conducted by our group, we described the construction of a *ciuA* gene mutant strain (Cp13), generated by using the *in vivo* insertional mutagenesis of the reporter transposon-based system TnFuZ in the T1 strain of *C. pseudotuberculosis*. The disruption of the *ciuA* gene led to a decrease in virulence and intracellular viability of the mutant strain. In addition, immunization of mice with the Cp13 strain elicited significantly IgG titers and higher survival rate (80%). The *ciuA* gene codes for a surface protein which is associated to a siderophore-base iron acquisition system; however, its role in the acquisition of iron in *C. pseudotuberculosis* is still largely unknown. Iron is crucial for the growth and development of many bacterial pathogens; however, within the host, the concentration that is freely accessible is much lower than the concentration required for bacterial growth. In intracellular bacteria, like *C. pseudotuberculosis*, iron homeostasis is transcriptionally controlled by the availability of iron in the environment, which allows the bacteria to counteract the limitation of the metal imposed by the host. Herein, our primary goal was to comprehensively characterize the transcriptional response of *C. pseudotuberculosis* under iron restriction. Differential gene expression analyses of the wild-type strain (T1 strain) and the (*ciuA*) Cp13 mutant strain (Cp13 strain) were conducted in low iron (LI) cultures, supplemented with the iron-chelator 2,2'-dipyridyl-DIP, in relation to high iron (HI) cultures in both strains. In this context, the data presented here shows a significant reduction in cellular density and the growth rate of the Cp13 mutant strain in relation to the parental strain under iron restriction. In total, 120 and 86 differentially expressed genes (DEGs) were identified in the T1 and Cp13 strains in response to iron restriction, respectively. 36 of these genes were identified in both strains and 26 of these genes had a similar expression pattern in both strains. Two transcription factors, known to be regulated by DtxR, were also identified between the overlapped DEGs of both strains, indicating a complex regulatory

network controlled by the availability of iron in *C. pseudotuberculosis*. Out of the 26 genes, 7 DEGs encode high-affinity hemin-binding proteins. Interestingly, 4 hemin-binding genes were identified located within genomic islands harboring known virulence factors, corroborating with the association of iron acquisition and virulence in pathogenic bacteria. We have also observed that the intracellular restriction of iron usage is key to the iron starvation response of *C. pseudotuberculosis.* Significant differences between mutant and parental strains were also observed in relation to the strain-specific DEGs, indicating divergence in the adaptive responses of the mutant and parental strains in response to iron restriction. Altogether, these findings provide a comprehensive analysis of the regulatory network and transcriptional response of the Cp13 mutant and T1 parental strain of *C. pseudotuberculosis,* adding relevant insights into the transcriptional adaptation of this important pathogen within an iron-restricted environment.

Keywords: *Corynebacterium pseudotuberculosis*, iron homeostasis, iron-regulated transcriptional factors, heme acquisition, differential gene expression

VI. DELINEAMENTO

Durante o período de desenvolvimento desta tese, eu participei diretamente da elaboração do projeto voltado para identificação de genes regulados pela disponibilidade de ferro o que gerou o manuscrito principal apresentado aqui. O manuscrito foi recentemente submetido ao periódico BMC GENOMICS.

De forma paralela, eu tive a oportunidade de participar na validação dos dados de RNA-seq obtidos para linhagem 258 de *C. pseudotuberculosis* em relação as alterações de osmolaridade e acidez utilizando PCR em tempo real. A linhagem 258, isolada de um cavalo infectado na Bélgica, foi o primeiro representante do biovar *equi* sequenciado pelo nosso grupo, sendo o genoma completo montado, anotado e depositado em maio 2009. O genoma sequenciado pela plataforma SOLiD possui um conteúdo de G+C de 52,15%, 12 rRNAs, 49 tRNAs, 2195 genes e 46 pseudogenes. A linhagem foi novamente sequenciada utilizando a plataforma Ion Torrent PGM e montada utilizando um sistema de alta resolução. Os dados para análise do perfil transcricional da linhagem 258 foram obtidos por Pinto em 2011, utilizando a plataforma SOLiD™. O trabalho teve como intuito, ampliar o conhecimento dos processos de adaptação bacteriana em resposta a condições que simulam o ambiente encontrado pela bactéria no hospedeiro durante o processo de infecção. Minha participação nesse trabalho, foi validar os dados de expressão por PCR em tempo real, com objetivo de concluir as análises do perfil transcricional da linhagem 258 de *C. pseudotuberculosis* permitindo a caracterização da genômica funcional em resposta a duas condições de estresse: um meio hiperosmótico contendo 2M de NaCl; e um meio ácido com pH 5, suplementado com HCl. O artigo foi submetido e aceito para publicação no periódico GENE (ISSN: 0378-1119) em Agosto de 2018 - *Transcriptome anylsis of Corynebacerium pseudotuberculosis biovar Equi in two condition of environmental stress* (doi: 10.1016/j.gene.2018.08.028). Nesse trabalho, eu e a Dra. Anne Cybelle, dividimos os créditos de primeira autoria. O trabalho foi adicionado ao anexo II dessa tese, como sugestão da banca de qualificação, para não entrar em conflito com o artigo principal sobre a influência do ferro na expressão gênica.

1. INTRODUÇÃO

1.1. *Corynebacterium pseudotuberculosis*

Corynebacterium pseudotuberculosis é uma bactéria Gram-positiva aeróbica facultativa, imóvel, pleomórfica, que não esporula e não forma cápsula (DORELLA *et al.*, 2006). É um patógeno intracelular facultativo, pertencente a classe Actinobacteria ao lado de outras espécies como Mycobacterium, Nocardia e Rhodococcus constituindo o grupo CMNR. O grupo apresenta características marcantes principais como conteúdo de bases nitrogenadas guanina e citosina em seu genoma variando entre 47-74% e parede celular composta principalmente de peptideoglicanos, arabinogalactano e ácidos micólicos (COELHO, 2007).

C. pseudotuberculosis é bioquimicamente classificada em dois biovares, *ovis* e *equi*, definidos assim pela presença/ausência da enzima nitrato-redutase, que permite a conversão do nitrato para nitrito. O biovar *ovis*, nitrato negativo, é mais comumente isolado de pequenos ruminantes, como cabras e ovelhas. Em contrapartida, o biovar *equi*, nitrato positivo, acomete preferencialmente os equinos e bovinos. É importante citar que os bovinos podem ser infectados pelos dois biovares, com predomínio do biovar *equi* (COSTA, 2010).

A bactéria apresenta grande relevância na medicina veterinária por ser o agente etiológico de uma variedade de doenças crônicas supurativas em diversos animais, como cabras, ovelhas, cavalos, alpacas, lhamas, búfalos, camelos e antílopes (BAIRD; FONTAINE, 2007a; KARIUKI; POULTON, 1982). Contudo, devido à grande importância para a agropecuária mundial, a linfangite ulcerativa (LU) em bovinos e equinos e a linfadenite caseosa (LC) em pequenos ruminantes, se destacam em relação ao considerável impacto econômico causado pelo o acometimento desses animais. Em termos gerais, essas doenças reduzem a produção e produtividade, comprometendo a pele, leite, carne, couro, a reprodução e a viabilidade dos animais (ARSENAULT *et al.*, 2003; BAIRD; FONTAINE, 2007a). Em termos econômicos, comprometem as exportações, aumentam os custos e diminuem consideravelmente os lucros (COSTA, 2010; FACCIOLI-MARTINS, P.Y.; ALVES, F.S.F.; PINHEIRO, 2014).

A LU é uma doença crônica que acomete vasos linfáticos cutâneos e resulta na presença de nódulos, abscessos subcutâneos e úlceras localizadas na extremidade inferior dos membros do animal (ALEMAN *et al.*, 1996). As úlceras contêm restos celulares necrosados com coloração escura e intenso tecido de granulação (MOTTA, R.G.; CREMASCO, 2010). Esses abcessos, têm uma evolução crônica e apresentam consistência dura e fistulam, deixando verter um exsudato sanguinolento ou purulento (amarelado) com estrias de sangue. Frequentemente o processo evolui para a formação de lesões ulcerosas. Em bovinos esses processos estão localizados geralmente na cabeça, pescoço, ombros, flancos e membros (MOTTA, R.G.; CREMASCO, 2010). Nos equinos os abcessos se restringem a região peitoral e ventral, sendo também conhecidos como a "febre do pombo". Os equinos infectados podem apresentar abscessos internos no pulmão, pericárdio, peritônio, rins, útero, mesentério, diafragma e outros, seguidos de febre, perda de peso e óbito (ALEMAN *et al.*, 1996; DOHERR *et al.*, 1998). As formas internas são raras entre os bovinos e se caracterizam pela presença de abscessos no trato respiratório e nos linfonodos internos (MOTTA, R.G.; CREMASCO, 2010).

Por sua vez, a LC é caracterizada pela formação de piogranulomas nos linfonodos superficiais, e, ocasionalmente, pulmões, baço, rins, fígado e sistema nervoso central (WILLIAMSON, 2001). Os piogranulomas contém pus de cor verde pálido a amarelo cremoso, inicialmente semilíquido, que posteriormente se espessam adquirindo a consistência caseosa, envolvido por uma cápsula de tecido conjuntivo. Em geral, os abscessos superficiais apresentam baixa morbidade, com reduzido impacto na saúde dos animais, enquanto que a presença de abscessos profundos e/ou abscesso pulmonares estão associados aos quadros com maior morbidade associados ao emagrecimento progressivo (caquexia) dos animais (ALVES, *et al*., 2007). A LC possui ampla distribuição mundial e tem sido identificada em países onde existem grandes rebanhos, como Austrália, Argentina, Nova Zelândia, África do Sul e Estados Unidos e em países da Comunidade Europeia (França, Itália, Grã-Bretanha, União Soviética), Chile, Uruguai, Canadá e Sudão (BAIRD; FONTAINE, 2007a). No Brasil, a epidemiologia da LC, apresenta prevalência clinica variável de 5% a 50%, sendo o nordeste brasileiro é a região mais afetada do país em virtude da maior concentração de rebanhos de caprinos e ovinos (MOTTA, R.G.; CREMASCO, 2010). Como descrito anteriormente, os prejuízos causados pela doença são muito grandes, impactando diretamente a produção de leite, lã, a depreciação da pele

do animal devido a presença de cicatrizes, a reprodução dos animais, condenação do couro e depreciação da carne destinada à exportação (COSTA, 2010).

A patogênese da LU ainda é pouco conhecida; contudo, já descrito na LC, observa-se que a formação de nódulos e abscessos se inicia pela inoculação bacteriana através de feridas localizadas na pele dos animais (BELCHIOR *et al.*, 2006). Uma vez inoculada, a bactéria é ativamente fagocitada por macrófagos, que migram para os linfonodos regionais. Nos linfonodos, a bactéria resiste ao compartimento fagolisossomal, mantendo-se viável no interior celular. A multiplicação bacteriana no interior dos macrófagos causa a lise e morte das células hospedeiras, acometendo o linfonodo drenante local e ocasionando a formação de lesões necróticas (PÉPIN *et al.*, 1994). Na tentativa de conter e eliminar o microrganismo, ocorre a formação de piogranulomas que podem coalescer e formar grandes abscessos. Microscopicamente a lesão se inicia pelo aparecimento de células epitelióides que precedem a formação de necrose caseosa. A massa caseosa é então circundada por células epiteliais e linfócitos, seguida por uma camada de tecido conjuntivo fibroso. Com a progressão da lesão, estas camadas celulares acabam por sofrer necrose, resultando no aspecto macroscópico desta patologia (JONES, T.C.; HUNT, R.D.; KING, 2000). A disseminação do organismo do linfonodo regional para outros órgãos e tecidos depende da virulência da linhagem, da carga bacteriana infectante e higidez do animal. Juntos estes fatores podem contribuir com a progressão da infecção e consequentemente para a forma visceral grave da doença. Contudo, é importante observar que a sobrevivência no interior dos fagócitos constitui uma etapa fundamental para o sucesso durante a disseminação e estabelecimento da bactéria no hospedeiro (JOLLY, 1966; PINTO, ANNE CYBELLE, 2011). Em parte, esse sucesso pode ser atribuído à presença de ácidos micólicos na parede celular bacteriana que possibilitam sua sobrevivência no interior das células hospedeiras, bem como a regulação de genes associados a virulência que tem papel fundamental na adesão, invasão, viabilidade, dispersão e patogênese de *C. pseudotuberculosis* (SCHUMANN, 2001).

A complexa estrutura de parede celular composta por uma camada externa de ácidos micólicos comum aos membros do grupo CMNR, é considerado um importante fator de virulência contribuindo para viabilidade bacteriana no interior de macrófagos (JANKUTE *et al.*, 2015). Como observado por Hard (1975), extratos de ácidos micólicos de *C. pseudotuberculosis* exercem um efeito negativo na atividade glicolítica, viabilidade

e integridade de membrana de macrófagos murinos e caprinos (HARD, 1975). Concordantemente, por mecanismos ainda desconhecidos, bactérias internalizadas por macrófagos se mantêm viáveis no interior celular, apesar dos macrófagos infectados apresentarem fagocitose e fusão lisossomal funcional (MCKEAN; DAVIES; MOORE, 2007b; TASHJIAN; CAMPBELL, 1983). Em *C. pseudotuberculosis*, esses lipídeos estão diretamente associados a manutenção da cronicidade da infecção, aos efeitos tóxicos e a sobrevivência da bactéria no ambiente intracelular, influenciando na patogenicidade da bactéria (BURKOVSKI, 2013).

Em adição ao fator de virulência citado, a regulação de genes de virulência ou genes associados a virulência, pode contribuir efetivamente para o estabelecimento da infecção no hospedeiro ou pode contribuir (genes associados à virulência) para a resistência bacteriana ao ambiente externo (fora do hospedeiro) e ao interno (dentro do hospedeiro) (PACHECO, 2010). Entre os genes mais estudados, são de particular importância para a virulência do gênero *Corynebacterium,* o gene *pld* e o cluster gênico formado pelo operon *fagABC* em associação ao gene *fagD* (MCKEAN; DAVIES; MOORE, 2007a). O gene *pld* codifica a proteína fosfolipase D (PLD), responsável por catalisar a dissociação da esfingomielina (um importante fosfolipídio da membrana citoplasmática dos vasos), em fosfato de ceramida e colina (SELVY *et al.*, 2011). O fosfato de ceramida associado a membrana compromete as células do epitélio vascular do hospedeiro, aumentando a permeabilidade e favorecendo a disseminação da bactéria a partir do local inicial da infecção. A fosfolipase D também apresenta atividade dermonecrosante e tóxica (MCNAMARA; CUEVAS; SONGER, 1995). A indução do gene *pld* após a infecção em macrófagos caprinos e roedores resulta em vacuolização do citoplasma, dano mitocondrial, dilatação do retículo endoplasmático. Já a expressão do gene *pld* atua de modo expressivo na viabilidade bacteriana, contribuindo na adaptação do microrganismo às mudanças ambientais durante a infecção e na manutenção da doença ativa (HODGSON *et al.*, 1992; MCNAMARA; BRADLEY; SONGER, 1994).

Em *C. pseudotuberculosis*, Billington *et al*. (2002) descreveu um operon (*fagABC*) com 32-47% de identidade com proteínas envolvidas no sistema de captação de ferro do tipo ABC em outras bactérias intracelulares. O cluster é composto pelo operon *fagABC* e pelo gene *fagD*, sendo os genes encontrados localizados dentro de uma ilha de patogenicidade juntamente com o gene *pld* no genoma de *C. pseudotuberculosis*. Neste mesmo trabalho,

mutantes *fagB(C)* mostraram uma redução significativa na capacidade de induzir abscessos em caprinos, corroborando a associação entre virulência e sistema de captação de ferro em bactérias patogênicas. A aquisição de ferro é um importante fator limitante na sobrevivência de bactérias no hospedeiro durante o processo de infecção, sendo considerada por muitos autores um dos principais fatores de virulência em bactérias patogênicas (FUKUSHIMA; ALLRED; RAYMOND, 2014; PARROW; FLEMING; MINNICK, 2013; SHELDON; LAAKSO; HEINRICHS, 2016). O ferro é o metal de transição mais abundante da crosta terrestre e um dos micronutrientes essenciais para a sobrevivência de diversos organismos. Em bactérias, o ferro age como cofator enzimático e desempenha papel importante em processos biológicos vitais, como replicação do DNA, transcrição e metabolismo (ANDREINI *et al.*, 2008). Muitos genes associados a obtenção de ferro se encontram localizados em ilhas de patogenicidade no genoma, sugerindo uma possível ligação entre sistemas de aquisição de ferro e virulência em bactérias patogênicas (BROWN; HOLDEN, 2002).

1.2. Homeostase de ferro em bactérias

Em condições fisiológicas, o ferro se encontra em dois estados de oxidação diferentes: a forma reduzida ferrosa Fe^{2+}, altamente solúvel, e a forma férrica Fe^{3+} insolúvel. O ferro Fe^{3+} é a forma mais comumente encontrada e está presente em altas concentrações de oxigênio. De forma oposta, a forma ferrosa é encontrada em ambientes com baixo pH e anóxia (ANDREWS; ROBINSON; RODRÍGUEZ-QUIÑONES, 2003). Apesar da sua importância em diversos processos, o ferro pode ser altamente tóxico, caso não ocorra uma regulação precisa das concentrações do metal no ambiente celular. Em altas concentrações, o ferro favorece a formação de radicais livres, onde a forma Fe^{2+} através da reação de Fenton (1), catalisa a formação do radical hidroxila (OH•) que é extremamente reativo, podendo lesar DNA, proteínas, açucares e lipídios. No estado metabólico normal, o superóxido favorece a oxidação de Fe^{2+} a Fe^{3+}. No entanto, se a concentração intracelular de superóxido é elevada ou na ausência de um substrato, o radical hidroxila formado pode oxidar outro íon Fe^{2+} perpetuando a reação de Fenton e aumentando assim, a concentração dos radicais hidroxila. A forma oxidada Fe^{3+} também pode reagir com moléculas de peroxido de hidrogênio formando radicais de hidroperoxila (HOO-) (AGUIAR *et al.*, 2007).

$$(1) Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + OH^{\bullet}$$

Devido a sua natureza reativa, a concentração de ferro livre nos fluidos extracelular do hospedeiro é quase nula, encontrando-se em sua maioria complexado a proteínas como transferrinas, lactoferrinas, ferritinas e proteínas contendo grupos heme (hemoglobina e mioglobina) (WEINBERG, 2000). O ferro complexado a proteínas minimiza os dados oxidativos do metal (formação de espécies reativas) e afeta diretamente a proliferação de patógenos, em virtude a sua baixa disponibilidade no ambiente. O controle da disponibilidade de ferro em mamíferos é conhecido como anemia imunomediada e se refere a anemia observada em indivíduos que apresentam processos inflamatórios, infecciosos ou neoplásicos crônicos (CASSAT; SKAAR, 2013). Com o intuito de superar esta restrição no hospedeiro, as bactérias empregam fatores transcricionais regulados pela disponibilidade do metal no ambiente que controlam a expressão de genes que codificam, por exemplo, para componentes do sistema de aquisição e transporte de ferro, controle do armazenamento e resistência oxidativa (ANDREWS; ROBINSON; RODRÍGUEZ-QUIÑONES, 2003; CARPENTER; WHITMIRE; MERRELL, 2009; SMITH, JAMES L, 2004). Além disso, inúmeras bactérias patogênicas utilizam os baixos níveis de ferro no hospedeiro como sinal para indução de genes de virulência.

1.2.1. Regulação da expressão gênica dependente de ferro em bactérias

Interessantemente, a homeostase de ferro em bactérias é controlada por fatores transcricionais específicos que respondem a disponibilidade do metal no ambiente, controlando a transcrição de genes que permitem a adaptação bacteriana à restrição de ferro imposta pelo hospedeiro (SKAAR, 2010). Em bactérias do gênero *Corynebacterium*, o processo de homeostase é controlado pela proteína reguladora DtxR (Diphtheria toxin repressor). O repressor da toxina diftérica -DtxR, foi incialmente descrito por regular a expressão da toxina diftérica em *C. diphtheriae* em resposta as concentrações ambientais de ferro (BULLEN, J.J. E GRIFFITHS, 1999). Contudo, são crescentes as evidências de que DtxR atua como um repressor ferro-dependente de diversos genes envolvidos em processos metabólicos, aquisição e armazenamento e virulência em inúmeras bactérias Gram-positivas (BRUNE *et al.*, 2006; D'AQUINO *et al.*, 2005; KUNKLE; SCHMITT, 2003; TROST *et al.*, 2010; WENNERHOLD; BOTT, 2006).

1.2.2. Mecanismos de captação e armazenamento de ferro em bactérias

Em *Corynebacterium,* a proteína repressora DtxR pode estar envolvida na regulação da síntese e captação de sideróforos em resposta à disponibilidade de ferro (KUNKLE; SCHMITT, 2005; TROST *et al.*, 2010). A utilização de sideróforos é uma das formas mais eficazes para a captação de ferro extracelular. Produzidos por bactérias e fungos, sideróforos são pequenas moléculas (<1 kDa) quelantes com alta afinidade ao ferro férrico. No hospedeiro, o ferro extracelular é encontrado ligado com alta afinidade a transferrina ou quelado a moléculas no plasma de baixa afinidade como albumina, citrato e aminoácidos (HOOD; SKAAR, 2012). Contudo, a constante de associação do ferro com os sideróforos bacterianos é maior (10^{50}) e supera a constante de associação do ferro a ferritina (10^{36}) (BULLEN, J.J. E GRIFFITHS, 1999). Sideróforos são sintetizados em condições de baixa disponibilidade de ferro e são secretados no meio extracelular onde se ligam ao Fe^{3+} facilitando a captação bacteriana (BRAUN, 2001). Tanto em bactérias Gram-positiva quanto em Gram-negativa, a produção de sideróforos é ferro dependente e está associada a virulência em uma variedade de patógenos como *S. aureus*, *Escherichia coli*, *Legionella pneumophila* e o *Bacillus anthracis* (CASSAT; SKAAR, 2013).

Em *C. pseudotuberculosis,* os clusters gênicos *fagABCD* e *ciuABCD-EF* codificam proteínas homólogas a sistemas de captação de ferro-sideróforos em outras bactérias (BILLINGTON; ESMAY; SONGER; JOST; *et al.*, 2002; KUNKLE; SCHMITT, 2005). O cluster *fagABCD*, descrito por Billington *et al.* (2002), é formado pelo operon *fagABC* e o gene *fagD,* sendo que os genes *fagA* e *fagB* codificam proteínas integrais de membrana que compõem uma permease transmembrana associada a subunidades proteicas com atividade ATPase codificada pelo gene *fagC,* formando o operon *fagABC*. O gene *fagD* está localizado a montante do operon *fagABC* e codifica uma lipoproteína acessória ancorada na superfície da membrana com domínio de ligação similar as proteínas de ligação a ferro-sideróforos (BILLINGTON; ESMAY; SONGER; JOST; *et al.*, 2002). Entretanto, apesar do envolvimento claro desses genes com a virulência de *C. pseudotuberculosis* e da regulação ferro dependente do cluster, os resultados gerados foram inconclusivos em relação a aquisição de ferro e produção de sideróforos em *C. pseudotuberculosis*.

Devido a sua importância nos processos celulares, a redundância na presença de sistemas de captação de ferro de mesma especificidade é comum no genoma bacteriano. Similar

ao operon *fagABC*, as proteínas codificadas pelo operon *ciuABCD* e o gene *ciuE* em *C. pseudotuberculosis* apresentam alta identidade com um sistema de captação de ferro-sideróforo em *C. diphtheriae*. (KUNKLE; SCHMITT, 2005). *Em diphtheriae*, o operon *ciuABCD* seria responsável por codificar as subunidades transportadoras do tipo ABC de ferro sideróforos, enquanto o gene *ciuE* estaria diretamente envolvido com a biossíntese de sideróforos. O cluster foi demonstrado ser regulado pela presença de ferro no meio; contudo, a biossíntese de sideróforos variou em relação as diferentes linhagens, tanto em relação aos mutantes *ciuA* quanto aos mutantes *ciuE*. Conclusivamente, as diferenças em ativação das regiões promotoras desses genes foram atribuídas a alguma especificidade relacionada a linhagem, ainda não conhecida (KUNKLE; SCHMITT, 2005).

Alternativamente, o ferro pode ser adquirido complexado ao grupo heme ou a hemoglobina. Em vertebrados, mais de 70% do ferro total encontra-se complexado ao grupo heme, associado a hemoglobina dentro de eritrócitos (TONG; GUO, 2009). O grupo heme é constituído por uma porção orgânica e um átomo de ferro (Fe^{2+}), sendo sua utilização direta ou degradação para liberação do metal, uma das formas encontradas pelos organismos patogênicos para compensar a baixa disponibilidade de ferro no hospedeiro (NOBLES; MARESSO, 2011). A utilização de heme parece ter uma relevância significativa na aquisição de ferro por bactérias patogênicas. Sistemas voltados para a aquisição da heme foram identificados no genoma de bactérias como *M. tuberculosis, C. diphtheriae*, *S. aureus*, *Bacillus anthracis, P. aeruginosa* e *Porphyromonas gingivalis* (KUNKLE; SCHMITT, 2003; MARESSO; GARUFI; SCHNEEWIND, 2008; MAZMANIAN *et al.*, 2003; OCHSNER; JOHNSON; VASIL, 2000; TULLIUS *et al.*, 2011).

O armazenamento de ferro em compartimentos intracelulares é extremamente relevante para a bactéria, tanto por gerar uma reserva interna que pode ser utilizada em condições de escassez como por proteger a bactéria dos danos oxidativos causados pela presença do íon metálico em solução (ANDREWS; ROBINSON; RODRÍGUEZ-QUIÑONES, 2003). Em bactérias, três tipos de proteínas armazenadoras de ferro são conhecidas: as ferritinas (Ftn); as bacterioferritinas (Brf) e as pequenas proteínas ligantes de DNA (Dps). Embora sejam evolutivamente distintas, essas proteínas apresentam sítios de ferro-oxidação conservados que permitem a oxidação de íons ferrosos em excesso e o armazenamento

de íons férricos no interior das proteínas como reserva interna bacteriana (HONARMAND EBRAHIMI; HAGEDOORN; HAGEN, 2015).

1.3. A pesquisa por determinantes de virulência em C. pseudotuberculosis

A pesquisa por determinantes moleculares de virulência em bactérias patogênicas tem dedicado atenção a genes transcricionalmente regulados em resposta as condições análogas as encontradas pela bactéria durante o processo de infecção. Durante o processo de infecção, patógenos como *C. pseudotuberculosis*, alternam entre ambientes com concentrações variáveis de oxigênio, osmolaridade, acidez e disponibilidade nutricional, do ponto inicial de infecção até a replicação intracelular em macrófagos (BECKER; SKAAR, 2014). Certamente, a habilidade de reconhecer e responder de forma efetiva as pequenas alterações químicas e físicas no ambiente assume um papel essencial na adaptação bacteriana a esses diversos nichos, regulando diretamente a expressão de genes que podem contribuir com a virulência do patógeno (PACHECO, 2010).

Com o intuito de ampliar o conhecimento dos genes envolvidos no processo de infecção de *C. pseudotuberculosis*, nosso grupo de pesquisa tem se dedicado aos estudos transcricionais relacionados a adaptação bacteriana às condições semelhantes as encontradas no hospedeiro. PINTO *et al.* (2014) submeteu a linhagem 1002 de *C. pseudotuberculosis* a diferentes estresses abióticos (pH, térmico e osmolaridade) e observou uma forte correlação entre a resposta ao estresse ácido e genes de virulência, além de caracterizar os diversos genes envolvidos na sobrevivência e no escape bacteriano no hospedeiro. CASTRO (2009) mostrou um aumento na expressão do fator regulador sigma (*sigA*) em reposta ao estresse oxidativo proveniente do peróxido de hidrogênio na linhagem 1002 de *C. pseudotuberculosis*. De forma paralela, PACHECO (2010) avaliou o perfil de expressão de proteínas secretadas na linhagem selvagem 1002 e na linhagem mutante *sigE*, demonstrando a importância do fator sigma na resistência bacteriana a condições de estresse ao ambiente intrafagossômico.

Paralelamente, estudos conduzidos por nosso grupo de trabalho no Laboratório de Genética Celular e Molecular visando a busca por determinantes gênicos de virulência e patogenicidade de *C. pseudotuberculosis*, levaram a identificação de uma promissora linhagem mutante, obtida através da técnica de mutagêneses aleatória utilizando a

linhagem selvagem T1 de *C. pseudotuberculosis*. O mutante recombinante, identificado como Cp13, foi obtido utilizando o sistema de transposição *in vivo* baseado no TnFuZ para identificação de genes codificadores de proteínas secretadas (DORELLA *et al.*, 2006). O sistema TnFuZ baseia-se na combinação do elemento transponível Tn4001 com o gene da fosfatase alcalina (*phopZ*) da bactéria gram-positiva *Enterococcus faecalis*. O gene reporter *phopZ* utilizado não apresenta os 18 aminoácidos iniciais que compõe sua sequência sinal, sendo sua atividade enzimática restabelecida pela inserção do transposon em loci gênicos que codificam proteínas secretadas, produzindo mutantes *phopZ* positivos. Esses mutantes são identificados visualmente pela degradação do substrato cromogênico 5-bromo-4-cloro-3-indolil fosfato pela proteína heteróloga secretada (Gibson e Caparon, 2002) (Figura 1).

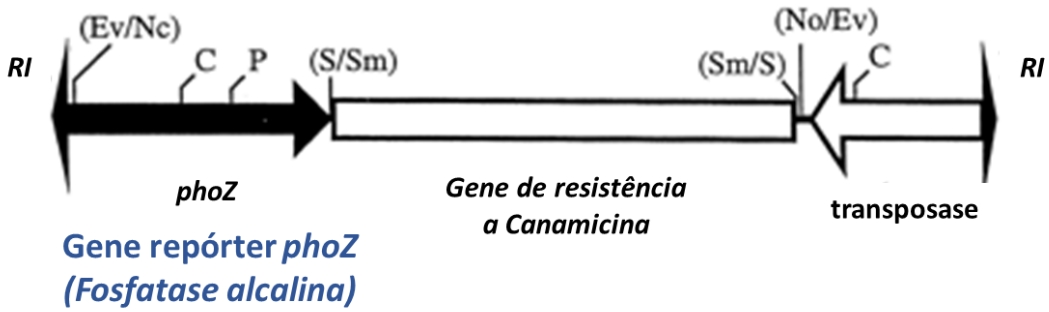


Figure 1. Estrutura do sistema TnFuZ

O sistema TnFuZ é delimitado por sequências de repetições invertidas (RI) em suas extremidades, gene da transposase e um gene de resistência a Canamicina. O gene da fosfatase alcalina (*phoZ*) de *Enterococcus faecalis* foi alterado, sendo removida a sequência codificadora do peptídeo sinal. A inserção deste transposon em loci gênicos que codificam proteínas exportadas leva a fusões que resultam em células secretando fosfatase e que são facilmente detectadas pela visualização in vitro do produto da degradação do substrato revelador 5-bromo-4-cloro-3- indolil fosfato (B.C.I.P.). Traduzido e adaptado de: (GIBSON; CAPARON, 2002).

O sequenciamento do DNA que flanqueia a região identificou que a inserção ocorreu no gene *ciuA* do operon *ciuABCD* que compõe um cluster gênico juntamente com os genes *ciuE* e *ciuF* (Figura 2).

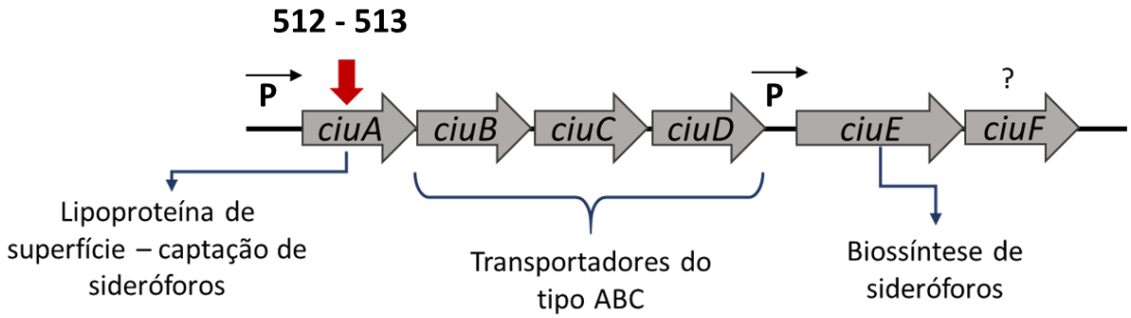


Figure 2. Representação esquemática do cluster *ciu* em *C. pseudotuberculosis*.
O gene *ciuA* está localizado em um operon juntamente com os genes *ciuB*, *ciuC* e *ciuD*. O operon faz parte de um cluster de 7 genes possivelmente envolvidos na aquisição e transporte de ferro-sideróforos em corynebacteria. O gene *ciuA* codifica uma lipoproteína de ligação ao ferro, o *ciuB-C* permeases de membrana do sistema ABC, o gene *ciuD* codifica uma ATPase de ligação e hidrólise de ATP e o gene *ciuE* codifica uma proteína relacionada a biossíntese de sideróforos. O papel do gene *ciuF* ainda não está bem esclarecido. A inserção do transposon TnFuZ ocorreu entre os nucleotídeos 512 e 513 do gene *ciuA*, interrompendo a sequência de 906 nucleotídeos do gene. A seta indica o ponto de inserção do transposon.

O operon c*iuABCD* codifica produtos que apresentam alta similaridade com proteínas relacionadas ao sistema ferro-sideróforo de transportadores identificado em *C. diphtheriae*, pertencentes a uma superfamília de transportadores do tipo ABC (ATP binding cassette) (KUNKLE; SCHMITT, 2005). Análises de homologia dos genes do operon *ciuABCD* em *Corynebacterium diphtheriae* demonstraram que, o gene *ciuA* codifica uma lipoproteína de ligação ao ferro, o *ciuB-C* permeases de membrana do sistema ABC e o gene *ciuD* codifica uma ATPase de ligação e hidrólise de ATP (Kunkle e Schmitt, 2005). O gene *ciuE* está localizado a jusante do operon *ciuABCD* e codifica uma proteína relacionada a biossíntese de sideróforos. KUNKLE e SCHMITT (2005) demonstraram o papel dos genes *ciuA* e *ciuE* em mutantes de *C. diphtheriae*, onde deleções do gene *ciuE* e *ciuA* resultaram na diminuição da atividade dos sideróforos e perda na habilidade de captação de ferro, respectivamente.

Apesar da relevância do gene *ciuA* no sistema de aquisição de ferro baseado em sideróforos ser pouco conhecida em *C. pseudotuberculosis*, trabalhos adicionais utilizando a linhagem mutante Cp13 demonstraram a importância desse sistema na viabilidade e virulência da espécie. Camundongos imunizados com a linhagem Cp13 apresentaram elevado nível de proteção (80%) quando desafiados com a linhagem virulenta MIC-6 de *C. pseudotuberculosis*, atingindo proteção superior a encontrada em

algumas vacinas comerciais em uso. Concordantemente, a linhagem Cp13 mutante induziu níveis significativos de anticorpos IgG e IgG2a em camundongos imunizados (DORELLA *et al.*, 2009). Mutantes Cp13 também apresentaram redução significativa na virulência e diminuição na viabilidade intracelular em ensaios de interação com macrófagos murinos (RIBEIRO *et al.*, 2014).

Como previamente abordado, a utilização de sideróforos de alta afinidade é uma das formas mais eficientes para a aquisição de ferro e, para muitas espécies patogênicas, a aquisição de ferro mediada por sideróforos é crucial para os processos de homeostase do metal. De forma geral, esses quelantes de ferro são induzidos em resposta à restrição do metal no meio, possibilitando a aquisição de ferro diretamente do ambiente extracelular do hospedeiro. Considerando a importância dessas moléculas entre as espécies patogênicas do grupo CMNR, consideramos plausível que o cluster *ciu*, mais especificamente o gene *ciuA*, assumiria a mesma importância em *C. pseudotuberculosis*. De forma mais direta, a hipótese do presente trabalho é que a interrupção do gene *ciuA* poderia estar associada a redução na sobrevivência e virulência da linhagem mutante, como observado por (DORELLA *et al.* (2006) e RIBEIRO *et al.* (2014). Portanto, diante de uma condição semelhante a encontrada no hospedeiro (restrição de ferro), nós poderíamos avaliar o papel dessa proteína e o envolvimento de sideróforos na adaptação de *C. pseudotuberculosis* comparando a resposta transcricional entre a linhagem selvagem e linhagem mutante. Contudo, tendo em vista escassez de informação envolvendo a expressão gênica de *C. pseudotuberculosis* em resposta à limitação de ferro, nosso objetivo principal foi caracterizar a resposta transcricional desse importante patógeno diante de uma condição que pudesse se assemelhara ao ambiente encontrado pela bactéria durante o processo de infecção.

2. JUSTIFICATIVA

Apesar da importância e abrangência das infecções causadas por *C. pseudotuberculosis*, ainda não foram identificados tratamentos e vacinas que atuem de forma eficiente no controle dessas infecções. De forma específica, a alta incidência de casos de LC em culturas de caprinos e ovinos no Brasil tem se tornado preocupante. Atualmente, o diagnóstico da LC é baseado na identificação bioquímica e molecular de material extraído de animais apresentando as manifestações clínicas da doença, excluindo a identificação de infecções subclínicas. Em relação ao tratamento, o uso de antibióticos é pouco eficaz, uma vez que são incapazes de penetrar na capsula dos abscessos (DORELLA *et al.*, 2009; DORELLA *et al.*, 2006). De modo agravante, as vacinas comerciais utilizadas no controle da LC apresentam eficiência variada em relação as diferentes espécies de animais, não oferecendo a proteção necessária contra a infecção (DORELLA *et al.*, 2006). Como mencionado anteriormente, a aquisição de ferro por patógenos bacterianos é essencial para o crescimento e o sucesso no estabelecimento da infecção. Tendo em vista sua importância, um grande esforço tem sido voltado para a compreensão de mecanismos de aquisição e homeostase de ferro que podem ser potencialmente utilizados no desenvolvimento de terapias no combate a infecções causadas por bactérias patogênicas. Dentro do hospedeiro, bactérias intracelulares como *C. pseudotuberculosis* são submetidas a uma condição de restrição férrica de modo a conter a proliferação e o avanço da infecção. Consequentemente, a concentração livre do ferro no hospedeiro é mantida em um nível inferior a concentração necessária para o crescimento bacteriano, encontrando-se complexado a proteínas de alta afinidade ou armazenado no interior das células do hospedeiro. Entretanto, diversos trabalhos têm demonstrado que a homeostase de ferro bacteriana é regulada em reposta à disponibilidade do metal no ambiente e que baixas concentrações de ferro, como encontrada no hospedeiro, podem induzir alterações significativas na expressão de genes envolvidos na aquisição, adaptação e virulência bacteriana. Embora os processos envolvidos na homeostase de ferro sejam bem caracterizados em espécies Gram-negativas, apenas nos últimos anos esses processos têm ganhado foco em bactérias Gram-positivas. Nessas espécies, genes que codificam proteínas voltadas para a aquisição de ferro na forma de heme ou complexados a sideróforos parecem desempenhar um papel crucial na aquisição de ferro por esses patógenos. No genoma de *C. pseudotuberculosis*, atualmente existem 20 genes

responsáveis por codificar proteínas homólogas a proteínas envolvidas no metabolismo de ferro de outras bactérias; contudo nossos conhecimentos sobre os mecanismos adaptativos e de regulação voltados para o metabolismo do metal em *C. pseudotuberculosis* ainda são bastante limitados. Neste contexto, a caracterização funcional do genoma de *C. pseudotuberculosis* em resposta à restrição de ferro pode contribuir de forma significativa no melhor entendimento dos mecanismos moleculares de homeostase de ferro por esse patógeno frente a uma condição próxima ao ambiente encontrado no hospedeiro durante o processo de infecção. De uma forma prática, uma melhor compreensão desses mecanismos pode servir como base para estudos futuros voltados para o controle de infecções causadas por *C. pseudotuberculosis*.

3. OBJETIVOS

3.1. Objetivo Geral

Avaliar a expressão gênica diferencial da linhagem selvagem T1 e da linhagem Cp13 mutante de *C. pseudotuberculosis* em resposta a baixa disponibilidade de ferro no meio.

3.2. Objetivos Específicos

Identificar e caracterizar funcionalmente os genes diferencialmente expressos na linhagem selvagem T1 e na linhagem Cp13 mutante (Δ*ciuA*) em resposta à limitação de ferro.

Comparar qualitativamente os dados de expressão gênica entre a linhagem selvagem e mutante, a fim de avaliar o papel da proteína CiuA em C. *pseudotuberculosis*.

Identificar as principais redes regulatórias transcricionais envolvidas na resposta bacteriana à disponibilidade de ferro.

4. METODOLOGIA

4.1. Linhagens bacterianas e condições de cultivo

As linhagens tipo-selvagem T1 e Cp13 *ciuA* mutante de *C. pseudotuberculosis* foram obtidas no Laboratório de Genética Celular e Molecular (LGCM) da UFMG. A linhagem T1 foi previamente isolada de granulomas caseosos em caprinos infectados no Brasil (Dorella et al., 2006). A linhagem mutante Cp13 foi obtida em trabalhos anteriores do nosso grupo através da técnica de mutagênese aleatória da linhagem selvagem T1 de *C. pseudotuberculosis* utilizando o sistema de transposição *in vivo* baseado no TnFuZ para identificação de proteínas secretadas (Dorella et al., 2006). O sistema baseia-se em um elemento transponível (Tn4001) combinado ao gene da fosfatase alcalina (*phoZ*) de *Enterococcus faecalis*, cujas regiões do promotor e codificadora do peptídeo sinal foram removidas. Uma vez que a fosfatase alcalina está ativa somente quando secretada, a inserção deste transposon em loci gênicos que codificam proteínas exportadas resulta em células com atividade de fosfatase alcalina. Essas células podem ser facilmente detectadas pela visualização *in vitro* do produto da degradação do substrato revelador 5-bromo-4-cloro-3-indolil fosfato (B.C.I.P.) (Dorella, 2009). Ambas linhagens foram rotineiramente cultivas em caldo infusão cérebro-coração (BHI) 0.05% Tween 80 e incubadas a 37ºC. Para as culturas em meio sólido, 1,5% de ágar bacteriológico foi adicionado ao meio de cultura. O antibiótico Canamicina foi acrescentado ao meio de cultura da linhagem selvagem a 25 μg/ml.

4.2. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de *C. pseudotuberculosis* em meio (BHI) e quantificação da concentração de ferro no meio BHI

Rotineiramente, *C. pseudotuberculosis* é cultivada em meio BHI (Brain Heart Infusion). O meio é muito utilizado na recuperação de microrganismo fastidiosos, incluindo bactérias aeróbicas e anaeróbicas. Devido a exigências nutricionais, *C. pseudotuberculosis* cresce bem em meios enriquecidos (CAMERON, SWART, 1965). Este meio é muito rico; no entanto, não se sabe a exata composição do meio, principalmente em relação a concentração de ferro. Assim, os níveis de ferro foram determinados utilizando a espectrometria de absorção atômica com chama para a

determinação da concentração do íon no meio BHI liquido. A análise foi realizada no departamento de química do Laboratório de Absorção Atômica da UFMG. De forma simplificada, a técnica se baseia na medição indireta da luz absorvida de um analito, onde átomos no estado fundamental (estado de menos energia) são capazes de absorver energia luminosa de tal forma que a molécula passe para um estado eletrônico excitado. Esse processo de transição eletrônica, provoca o aparecimento de uma banda de absorção no espectro, que representa a transmitância da luz na amostra em função da energia, podendo assim realizar uma determinação quantitativa do analito presente (SKOOG, WEST, HOLLER, 2005).

A fim de avaliar o impacto da interrupção do gene *ciuA* no crescimento da linhagem Cp13 mutante, curvas de crescimento foram realizadas com as linhagens tipo selvagem (T1) e mutante (Cp13) em meio BHI (Figura 3). O*vernight* pré-culturas (~20-24 hrs) das duas linhagens foram diluídas a uma densidade ótica inicial de 0.02 (DO_{600} = 0.02) em meio BHI fresco e incubadas por 10 horas sob constante agitação a 37ºC. O monitoramento do crescimento celular das culturas foi realizado a cada hora utilizando medidas de densidade óptica a 600nm. O impacto da interrupção do gene *ciuA* na linhagem mutante também foi avaliado através de análises de viabilidade bacterina por plaqueamento e contagem de unidades formadoras de colônia (UFC), como padronizado previamente no nosso grupo (Castro, 2009). A viabilidade celular foi avaliada a partir das culturas incubadas por 10 horas para as duas linhagens. Curvas de crescimento e análise estatística da viabilidade das linhagens tipo-selvagem e mutante de *C. pseudotuberculosis* foram realizadas utilizando o software GraphPad Prism (GraphPad Software, Inc).

4.3. Preparo do ligante de ferro 2,2-bipiridina (BIP) e determinação da concentração de uso

Com a finalidade de determinar a concentração mais adequada do ligante para utilização nos ensaios de depleção de ferro, avaliou-se a redução do crescimento das linhagens T1 e Cp13 em meio BHI suplementado com diferentes concentrações do quelante em relação ao meio BHI não suplementado. A redução do ferro livre no meio BHI foi obtida adicionando o quelante de ferro 2,2-bipiridina (BIP) e as amostras cultivadas suplementadas com esse quelante foram identificadas como LI, do inglês *low iron*. Ao longo do texto, o meio BHI não suplementado foi identificado como HI, do inglês *high iron*, para facilitar a comparação entre as culturas suplementadas com o quelante

bipiridina (LI) e não suplementadas com o agente quelante (HI). A bipiridina faz parte dos compostos aromáticos heterocíclicos, que são assim classificados por apresentarem um elemento diferente do carbono em seu anel aromático. A 2,2-bipiridina é obtida a partir da dimerização da piridina, sendo considerado um importante agente quelante de íons metálicos. Sua capacidade de formar complexos com sais de ferro tem sido amplamente estudada e seu alto potencial de oxidação tem sido empregado em diversos procedimentos analíticos como indicador de oxi-redução (KAES; KATZ; HOSSEINI, 2000). O ligante 2,2-bipiridina se liga predominantemente a moléculas de Fe^{2+} formando complexos 3:1 com átomos de ferro, inibindo assim sua disponibilidade e afetando de forma significativa o crescimento de diversos tipos celulares. Além disso, foi demonstrado que a bipiridina atravessa a membrana citoplasmática quelando diretamente o ferro no citosol, induzindo uma resposta celular de depleção ao metal nas células hospedeiras e em bactérias intracelulares (KALINOWSKI, 2005). A bipiridina é um dímero sólido cristalino de forma molecular $(C5H4N)_2$, pouco polar e insolúvel em água (KAES; KATZ; HOSSEINI, 2000). O*vernight* pré-culturas (~20-24 hrs) das duas linhagens foram diluídas a uma densidade ótica inicial de 0.02 (DO_{600} = 0.02) em meio BHI fresco suplementado com 200-220-240-260-280 μM da bipiridina e incubadas por 14 horas sob constante agitação a 37ºC. Culturas controle da linhagem T1 e Cp13 em meio BHI não suplementado com o agente quelante foram conduzidas paralelamente as culturas suplementadas para comparação dos resultados. O crescimento das culturas foi monitorado a cada hora por medidas de densidade óptica a 600 nm e as curvas de crescimento foram plotadas utilizando o software GraphPad Prism (GraphPad Software, Inc).

4.4. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de *C. pseudotuberculosis* em meio (BHI) suplementado com 250 μM do ligante 2,2-bipiridina

A fim de avaliar o crescimento e viabilidade das linhagens de *C. pseudotuberculosis* à restrição de ferro, o*vernight* pré-culturas (~20-24 hrs) das duas linhagens foram diluídas a uma densidade ótica inicial de 0.02 (DO_{600} = 0.02) em meio BHI fresco suplementado com 250 μM da bipiridina e meio BHI não suplementado e incubadas por 10 horas sob constante agitação a 37ºC. O monitoramento do crescimento celular das culturas foi realizado utilizando medidas de densidade óptica a 600nm a cada hora e a viabilidade celular foi avaliada pela contagem de unidades formadoras de colônias a cada duas horas

para as duas linhagens. Curvas de crescimento e análise estatística da viabilidade das linhagens tipo-selvagem e mutante de *C. pseudotuberculosis* foram realizadas utilizando o software GraphPad Prism (GraphPad Software, Inc).

4.5. Desenho experimental para a análise de expressão diferencial em resposta à limitação de ferro

A figura 7 apresenta o desenho experimental adotado para o preparo das amostras e para as análises de expressão diferencial das linhagens T1 e Cp13. Os números em parênteses indicam as etapas experimentais. Como previamente descrito, a avaliação do perfil transcricional por RNA-seq foi utilizada para identificar alterações diferenciais na expressão gênica da linhagem parental T1 e da linhagem Cp13 mutante de *C. pseudotuberculosis* em relação a baixa disponibilidade de ferro no meio. A restrição de ferro no meio foi realizada suplementando culturas com 250 uM do agente quelante de ferro bipiridina. Essas culturas suplementadas foram identificadas como LI (*low iron*). Paralelamente, culturas individuais de cada linhagem foram incubadas no meio BHI controle, utilizado rotineiramente para o cultivo de *C. pseudotuberculosis*. Essas culturas identificadas como HI (*high iron*), apresentam concentrações indefinidas de nutrientes e alta concentração de ferro. O*vernight* pré-culturas (~20-24 hrs) das duas linhagens foram diluídas a uma densidade ótica inicial de 0.02 (DO_{600} = 0.02) em meio BHI fresco suplementado com 250 μM da bipiridina e meio BHI não suplementado e incubadas por 6h30 sob constante agitação a 37ºC. Para monitorar o crescimento das linhagens nas diferentes condições, medidas de densidade óptica a 600 nm das culturas foram realizadas a cada hora. Os valores de densidade ótica foram utilizados para estimar a taxa de crescimento para cada condição de acordo com a seguinte fórmula:

$$Taxa\ de\ crescimento\ (Gr) = \log DO_f - \log DO_i/(t_f - t_i),$$

Onde, Gr = growth rate ou taxa de crescimento é igual ao logaritmo na base 10 da DO final menos o logaritmo na base 10 da DO inicial em função do tempo (Tempo final menos tempo inicial).

A resistência das linhagens de *C. pseudotuberculosis* às diferentes condições de estresse *in vitro* também foi avaliada através de análises de viabilidade bacteriana por plaqueamento e contagem de unidades formadoras de colônia (UFC). Finalizando o

período de incubação (6h30) alíquotas de ambas as linhagens provenientes de culturas suplementadas (LI) e não suplementas (HI) foram centrifugadas por 10 minutos a 5000 rpm (4˚ C). O sobrenadante celular foi descartado e o precipitado ressuspendido em 300 µL de tampo de lise Tris/EDTA/SDS (Tris HCl pH8 50 mM/L–EDTA pH8 10 mM/L) para cada 10 mL de cultura utilizada. A suspensão celular foi homogeneizada com 0,1% de SDS e 1mL de TRIzol® Reagent (Invitrogen™) foi acrescentado. (3) O RNA total de cada amostra foi extraído utilizando o reagente Trizol, sendo seguidas as recomendações do fabricante com as seguintes adaptações. Após adição do Trizol 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 6.200 rpm por 2 ciclos de 15 segundos, com intervalo de 5 segundos entre estes. As amostras foram então centrifugadas por 1 min e o sobrenadante dos tubos transferido para novo tudo de 2 mL e incubadas em banho seco a 60° C por 15 min seguindo todo protocolo original. Com o objetivo de remover traços de DNA genômico presentes nas amostras extraídas, o RNA total foi tratado utilizando o kit Ambion® TURBO DNA-free™ DNase (Invitrogen), segundo recomendações do fabricante. A eluição foi feita com 50 µL de água mili-Q RNase-Freee e a quantificação do RNA total presente em cada amostra foi avaliada pela razão Abs_{260}/Abs_{280} obtida utilizando espectrofotometria (NanoDrop, Thermo Scientific). A qualidade do RNA total extraído foi também avaliada através do perfil eletroforético das subunidades de rRNA obtido pelo equipamento Agilant 2100 Bioanalyzer. As amostras de RNA com razão de absorbância (Abs_{260}/Abs_{280}) superior a 1.9 foram consideradas adequadas para o sequenciamento e foram utilizadas para o preparado das bibliotecas de cDNA. Cerca de 5 µg de RNA total de cada amostra foram submetidos a depleção do rRNA utilizando o *kit* da Invitrogen Ribominus™ Transcriptome Isolation (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 depletado foi utilizado para o preparo das bibliotecas de cDNA utilizando o protocolo padrão do kit Ion Total RNA-seq v2 for Whole Transcriptome Library (Thermo Scientific). As bibliotecas amplificadas de cDNA foram sequenciadas utilizando a plataforma Ion Proton™ e todo o procedimento foi realizado no departamento de Bioquímica em parceria com a Universidade Federal do Paraná.

4.6. Estatísticas do sequenciamento e avaliação da qualidade das reads

A obtenção dos dados brutos gerados pelo sistema Ion Proton foi realizada com o software Torrent Suite 5.0.5 (Thermo Scientific) e a qualidade das *reads* sequenciadas foi avaliada utilizando o programa FastQC (http://www.bioinformatics.babraham.ac.uk/publications.html). As *reads* foram processadas utilizando a opção SLIDINGWINDOW do software Trimmomatic v0.36 (Bolger et al., 2014) para remoção de dados que apresentam um valor *phred* de qualidade inferior a 10 em relação ao valor médio de qualidade da janela de leitura. As *reads* processadas foram mapeadas ao genoma da linhagem 1002B de *Corynebacterium pseudotuberculosis* (NZ_CP012837.1) utilizando o Torrent Mapping Alignment Program (TMAP) v3.4.1. O TMAP alinha *reads* de tamanho variável e é otimizado para lidar com os erros de homopolímeros gerados pelo Ion Torrent. Para o alinhamento, foi utilizado o comando -mapall que combina os 4 algoritmos disponíveis (map1, map2, map3 e map4), com penalidade de 3 para *mismatch* e pontuação de 1 para cada match, default do programa. Os arquivos gerados pelo alinhamento foram manipulados utilizando o pacote de ferramentas do SAMTOOLS v.1.5. Esses arquivos foram então ordenados, combinados e indexados. Dados estatísticos simples sobre o mapeamento foram produzidos com a ferramenta Flagstat. Os arquivos ordenados foram utilizados como arquivos de entrada para o programa HTSeq-count, pacote do software HTSeq, que gera uma matriz de contagem dos transcritos mapeados a éxons únicos no genoma de referência. (5) A expressão gênica dos dados referentes a contagem de reads mapeadas a genes da referência foi utilizada para a quantificação relativa da expressão das amostras cultivadas em meio LI em relação as amostras cultivadas em meio HI. A identificação dos genes diferencialmente expressos (GDEs) em resposta a restrição de ferro foi realizada utilizando o DESeq2/Bioconductor. A primeira etapa do programa é a normalização dos dados brutos de contagem em relação ao total de sequências obtidas e em relação a composição do RNA por biblioteca, o que possibilita a subsequente comparação entre as diferentes bibliotecas. Uma das características do DESeq2 é a utilização de um modelo linear generalizado com distribuição binomial negativa para estimar a dispersão e contagem média de transcritos alinhados para cada gene entre as condições experimentais. O valor de *p* é calculado utilizando o Teste Wald e assumindo que a hipótese nula é verdadeira, ou seja, não existe diferença de expressão entre as duas condições experimentais, DESeq2 calcula a probabilidade *p* (*valor p*) da diferença observada entre as condições para rejeição da hipótese nula. Por último, calcula-se para cada gene o valor *p* ajustado para múltiplos testes utilizando o método de Benjamini-

Hochberg (Benjamini e Hochberg, 1995), onde a taxa de erro global é controlada ajustando o False Discovery Rate (FDR). O pacote DESeq2 informa também o valor da razão de expressão (fold-change) em logaritmo de base 2 (log2 fold-change) de uma condição em relação à outra. Especificamente, o gene é considerado diferencialmente expresso, se o valor de fold-change exceder um valor de corte pré-estabelecido. Assim, foram considerados diferencialmente expressos apenas genes com valor *p-ajustado* igual ou inferior a 0.05 e cujo valor da razão de expressão (fold-change) foi igual ou superior a 1.5 vezes. A expressão relativa é apresentada como a razão log2fold da expressão LI versus HI, detalhando os genes diferencialmente expressos identificados em resposta restrição de ferro. A análise ontológica funcional dos genes diferencialmente expressos em resposta a restrição por ferro foi realizada utilizando a plataforma online GOfeat. Adicionalmente, análises de interações proteína-proteína (PPI) foram realizadas para investigar associações funcionais e diretas entre os produtos proteicos dos GDEs. PPIs e enriquecimento funcional foram construídos utilizando a ferramenta online STRING. Gipsy (Genomic Island prediction software) foi utilizado para a predição de ilhas genômicas no genoma de *C. pseudotuberculosis* em relação a espécie não patogênica *C. glutamicum* (NC_003450). Uma análise qualitativa comparando os resultados de expressão diferencial obtidos na linhagem Cp13 mutante em relação a sua linhagem parental T1 foi conduzida com o objetivo de investigar o papel do gene *ciuA* na adaptação de *C. pseudotuberculosis* à restrição de ferro.

5. RESULTADOS

5.1. Artigo: Transcriptome Profile of *Corynebacterium pseudotuberculosis* in Response to Iron Limitation (Submetido para publicação no periódico BMC Genomics – Online ISSN: 1471-2164).

O material suplementar referente a esse artigo se encontro no Anexo I:

Additional file 1: Table S1. Sequencing statistics ……112

Additional file 1: Table S2. Total number of processed reads from High Iron (HI) and Low Iron (LI) experimental condition……113

Additional file 1: Table S3. Mapping results…….114

Additional file 1: Table S4. Raw gene counts…….115

Additional file 1: Figure S1. Sample to sample principal component analysis (PCA) ……...116

Additional file 2: Figure S2. *C. pseudotuberculosis* growth rate ……...117

Additional file 2: Figure S3. Effects of ethanol on bacterial growth……...118

Additional file 3: Table S5. Differentially expressed genes identified in the T1 strain in response to iron limitation……...119-123

Additional file 3: Table S6. Differentially expressed genes identified in the Cp13 mutant in response to iron limitation……...124-127

Additional file 4: Table S7. Table of gene ontology terms assigned to DEGs identified in the T1 strain……128-130

Additional file 4: Table S8. Table of gene ontology terms assigned to DEGs identified in the Cp13 mutant…….131-132

Additional file 5: Figure S4. Alignment of the amino acid sequence of the HtaA domain from protein products of the upregulated genes *htaC*, *htaA*, *htaF*, *htaG* and *Cp_3070* genes of *C. pseudotuberculosis*……133

Additional file 5: Figure S5. Structural protein domain characteristics of the HtaA, HtaC, HtaF, HtaG and Cp_3070 proteins……134

Additional file 5: Figure S6. Alignment of the amino acid sequences of the HmuU protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)……135

Additional file 5: Figure S7. Alignment of the amino acid sequences of the HmuT protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)……136

Additional file 5: Figure S8. Alignment of the amino acid sequences of the DNA-binding regulator *hrrA* of the two-component regulatory system *hrrSA* between *C. pseudotuberculosis* (Cp), *C. glutamicum* (Cg) and *C. diphtheriae* (Cd)…….137

Additional file 5: Table S9. Genomic Islands prediction……138

5.1.1. Abstract

Background: Iron is an essential micronutrient for the growth and development of virtually all living organisms, playing a pivotal role in the proliferative capability of many bacterial pathogens. The impact that the bioavailability of iron has on the transcriptional responses of bacterial species in the CMNR group has been widely reported for some members of the group, but it hasn't yet been deeply explored in *Corynebacterium*

pseudotuberculosis. Here we describe for the first time a comprehensive RNA-seq whole transcriptome analysis of the T1 wild-type and the Cp13 mutant strains of *C. pseudotuberculosis* under iron restriction. The Cp13 mutant strain was generated by transposition mutagenesis of the *ciuA* gene, which encodes a surface siderophore-binding protein involved in the acquisition of iron. Iron-regulated acquisition systems are crucial for the pathogenesis of bacteria and are relevant targets to the design of new effective therapeutic approaches.

Results: Transcriptome analyses showed differential expression in 106 genes within the parental T1 strain and 80 genes in Cp13 mutant under iron restriction. Twenty-five of these genes had similar expression patterns in both strains, including up-regulated genes homologous to the hemin uptake *hmu* locus and two distinct operons encoding proteins structurally like hemin and Hb-binding surface proteins of *C. diphtheriae*, which were remarkably expressed at higher levels in the Cp13 mutant than in the T1 wild-type strain. These hemin transport protein genes were found to be located within genomic islands associated with known virulent factors. Down-regulated genes encoding iron and heme-containing components of the respiratory chain (including *ctaCEF* and *qcrCAB* genes) and up-regulated known iron/DtxR-regulated transcription factors, namely *ripA* and *hrrA*, were also identified differentially expressed in both strains under iron restriction.

Conclusion: Based on our results, it can be deduced that the transcriptional response of *C. pseudotuberculosis* under iron restriction involves the control of intracellular utilization of iron and the up-regulation of hemin acquisition systems. These findings provide a comprehensive analysis of the transcriptional response of *C. pseudotuberculosis*, adding important understanding of the gene regulatory adaptation of this pathogen and revealing target genes that can aid the development of effective therapeutic strategies against this important pathogen.

Keywords: *Corynebacterium pseudotuberculosis*, iron homeostasis, iron-regulated transcriptional factors, heme acquisition, differential gene expression.

5.1.2. Background

Caseous lymphadenitis (CLA) is a chronic, debilitating infection caused by the Gram-positive bacterium *Corynebacterium pseudotuberculosis*. CLA affects mainly small ruminants, such as sheep and goats, and is the leading cause of economic loss associated with the reduction of wool, meat and milk production, and also results in carcass and skin condemnation in the majority of sheep and goat production areas [1]. CLA infection is typically initiated by inoculation via oral, respiratory, and membrane wounds [2]. Once inside, the bacteria are phagocytosed by macrophages, which are then drained to local lymph nodes. Once internalized, the bacteria evade the cellular immune response mechanisms, being able to survive and rapidly multiply within the macrophage phagosome. Bacterium proliferation results in the death of the host cell and the subsequent release of bacteria into the extracellular environment, which can then spread through the lymphatic system to regional lymph nodes and internal organs [3]. Clinical signs of the disease include the formation of caseous abscesses either in superficial or internal lymph nodes. While antibiotic therapy is not usually helpful, management of the disease combines the identification and removal of infected animals along with vaccination of healthy animals [2].

In recent years, novel mechanisms have been identified as potential targets to combat infections caused by corynebacteria. Considerable progress has been made in understanding mechanisms involved in iron acquisition and homeostasis in pathogenic species that can potentially be used in the development of strategies to control the pathogen [4]. Iron is fundamental in a variety of cellular functions and its capacity to directly participate in redox reactions, donating or accepting electrons, highlights its central importance in biological processes including DNA synthesis, enzyme and redox catalysis, electron transport, respiration, cellular survival and growth [5,6]. Given its absolute requirement by all bacteria, iron acquisition is a fundamental step for the establishment of the infection by pathogenic species, and iron withholding is one of the first host defense mechanism used to prevent bacterial growth [7]. Under physiological conditions, free ferric iron is insoluble and therefore is poorly accessible to support microbial growth. Within the host, iron is kept within intracellular proteins such as hemoglobin, cytochrome, ferritin or is found chelated by transferrin or lactoferrin in the extracellular compartment, in an attempt to further decrease iron accessibility to pathogens [8,9]. To overcome the low bioavailability of the metal imposed by the host,

pathogenic bacteria coordinate a complex and sensitive iron-dependent transcriptional regulatory network that controls the expression of genes related to iron homeostasis, including uptake, storage and iron-dependent systems according to the availability of the metal [10]. In pathogenic bacteria, the uptake is the most tightly regulated process in iron homeostasis, which is highlighted by the diverse complement of mechanisms used to efficiently exploit the host iron sources [10,11].

In a previous report, we described the construction of a *ciuA* mutant in *C. pseudotuberculosis*, designated Cp13 [16]. The Cp13 mutant was generated by using the in vivo insertional mutagenesis of the reporter transposon-based system TnFuZ in the strain T1 [17]. The molecular characterization of the Cp13 mutant showed that the insertion disrupted the *ciuA* gene, which encodes a putative-iron transport binding protein of the *ciuABCD* operon [16]. In *C. pseudotuberculosis*, the *ciuA* gene appears to be organized into the *ciuABCD* operon comprising a six gene-cluster, *ciuABCDEF*, which shares high protein similarity with components of siderophore ABC-type transport systems. The ciuA gene shares amino acid similarity to periplasmatic binding proteins from Gram-negative bacteria and 73% sequence identity to a putative iron transport binding protein from *C. diphtheriae* (DIP0582). In *C. diphtheriae*, the predicted product of the *ciuA* gene is a lipoprotein of 298 amino acids responsible for the transportation of iron-siderophore complexes. The *ciuB*, *ciuC* and *ciuD* genes encode proteins possibly related to siderophore ABC-type transporters. Downstream from the *ciuABCD* operon is the ciuE gene, whose predicted product shares homology to enzymes responsible for the biosynthesis of siderophore aerobactin [18,19]. The *ciuF* gene has low similarity to integral membrane efflux proteins and its role in siderophore transportation is still unknown. In addition to the *ciuABCDEF* genes, the genome of *C. pseudotuberculosis* harbors another putative iron siderophore system, formed by the *fagABCD* genes. Similar to ciu clusters, the *fagABC* operon and the *fagD* gene encode proteins with similarity to siderophore ABC-type transporters identified in other pathogenic species. It has also been previously demonstrated that the cluster is regulated by iron, being associated with the pathogenesis of *C. pseudotuberculosis* [20]. While the relevance of the ciu cluster regarding iron acquisition remained unclear for *C. pseudotuberculosis*, experimental studies conducted by our group with the Cp13 mutant strain have demonstrated a link between the disruption of the *ciuA* gene and virulence of the mutant strain. The Cp13 mutant presented a substantial reduction in virulence with a decrease in abscess formation

in experimental CLA infections [16]. Moreover, in a previously reported vaccination trial conducted with the Cp13 mutant, our group has demonstrated that the Cp13 mutant was able to elicit both humoral and cellular responses in immunized mice, with significant production of IgG and IgG2a. In addition, 80% of the mice immunized with the Cp13 mutant strain survived after being challenged with a virulent C. pseudotuberculosis strain. This mutant strain also showed reduced intracellular viability in murine cells [21].

As previously stated, iron acquisition is crucial for the growth and development of many bacterial pathogens; however, the concentration that is freely accessible within the host is much lower than the concentration required for bacterial growth [10]. While the iron uptake process has already been well characterized in Gram-negative bacteria, our knowledge regarding iron acquisition in Gram-positive pathogens has only recently been examined. In high G+C Gram-positive bacteria, genes encoding iron and heme acquisition proteins, and high-affinity iron scavenging siderophores associated with energy-dependent iron transporter systems seem to play a pivotal role in the procurement of iron by these pathogens [11–14]. These high-affinity iron acquisition systems are usually detectable under iron-restricted conditions and are regulated by iron-dependent transcription factors (TF) in response to the bioavailability of the metal [15]. In the *C. pseudotuberculosis* genome, there are twenty putative described genes directly involved in the acquisition and metabolism of iron; of which many have been associated with the pathogenesis of the bacterium [16,23,24]. The virulence of *C. pseudotuberculosis*, as for many pathogenic species, depends on their ability to survive the harsh iron-limited environment inside the host. However, despite the clear importance of iron during infection, our knowledge regarding *C. pseudotuberculosis* sensing and regulating of genes involved in iron homeostasis is limited. Driven by the paucity of information on the acquisition of iron in *C. pseudotuberculosis*, our primary goal was to comprehensively characterize the transcriptional iron response of this important pathogen, as well as its association with growth and virulence under a condition similar to the iron-limited environment which is encountered by the bacterium during the infection inside the host. In addition, since siderophore-mediated iron acquisition seems to play a critical role in iron homeostasis for many bacteria pathogens, we hypothesized that the ciuA gene was also essential for the uptake of iron in *C. pseudotuberculosis*. To address this hypothesis, we compared gene expression and growth profiles between the mutant (Cp13) and wild-type (T1) strains in the presence of the iron chelator 2,2'-dipyridyl-DIP (Low Iron) and

in the absence of the chelating agent (High Iron). To our knowledge, this is the first report addressing the complexity of the *C. pseudotuberculosis* response to iron limitation.

5.1.3. Results and Discussion

Sequencing statistics and data quality assessment

During infection, the expression levels of many iron homeostasis genes in the bacteria are regulated in direct response to the bioavailability of the metal being usually detectable under iron-restricted conditions [15]. To investigate the gene expression of *C. pseudotuberculosis* under iron restriction a high-throughput RNA sequencing (RNA-seq) transcriptome profiling was used to characterize the differential gene expression (DGE) of the wild-type T1 strain and the Cp13 mutant between low iron (LI) and high iron (HI) cultures (LI vs HI). Two rounds of RNA sequencing were performed on the Ion proton system yielding over 120 million single-end reads. Sequencing data were collected across 4 independent sets of paired biological samples (HI and LI) for the Cp13 mutant and 3 sets of paired samples for the wild-type T1 strain and labeled as batches A-C (See additional file 1: Table S1 for batch identification). One HI sample from the T1 strain had a low total number of sequenced reads, with less than <0.4% of the total sequences for its LI paired-sample, 38.405 and 12.560.458 respectively. As a result, both of these paired samples were removed from further analyses. Detailed information on the sequencing statistics of each sample is provided in the Additional file 1: Table S1.

The remaining samples were quality checked and had an average read quality Phred score of 24-25 with a mean per base quality score of over 20. An average base quality trimming was performed, with less than 1% of the total reads from each set excluded (Additional file 1: Table S2). Trimmed reads were mapped to the 1002B reference genome. The percentage of reads mapping to the genome ranged from 72.6 to 98.9% (Additional file 1: Table S3). Unambiguously aligned reads were identified by HTSeq-count, which excluded all reads mapped to or overlapping with more than one gene in the reference (Additional file 1: Table S4).

Principal component analysis (PCA) was used to assess the relationship between samples for both strains (Additional file 1: Figure S1). PC1 and PC2 explain 85% of the total variance in the dataset. PC1 revealed a clear separation between experimental conditions

(HI and LI). Except for one HI sample, batches B and C of the Cp13 mutant and T1 strain datasets grouped together, indicating low inter-strain variability among replicates of the same experimental batch. PC2 clearly evidenced the variability of the paired-sample 1 of the Cp13 mutant (batch A) batches B-C of the same strain. Batch effect was attributed to sample processing and batch was added as a variable into the experimental design for subsequent differential expression analysis of the Cp13 mutant (Additional file 1: Figure S1).

Effects of iron chelating on viability and growth of *C. pseudotuberculosis* strains

To assess the effects of iron depletion on bacterial viability and growth we compared the number of CFU/mL and the rate of proliferation for each strain between a low (LI - dipyridyl-chelated) and high (HI - non-chelated) iron conditions (Fig. 3). Both cultures showed an equivalent ratio in bacterial growth, with no statistical differences observed in the rate of growth in the high iron medium (student t-test with $p > 0.05$), with the Cp13 mutant strain achieving cellular density similar to the parental T1 strain (Additional file 2: Figure S2). Based on what is known regarding iron acquisition in corynebacteria, it is likely that the presence of redundant iron uptake systems in the chromosome of these species could be responsible for the acquisition of iron distinctively from the ciu system in an iron rich environment. As expected, both bacterial cultures were iron-stressed in the iron depleted environment with a statistically significant decrease in bacterial growth in the Cp13 mutant in relation to the parental T1 strain (p-value=0.0016) (Additional file 2: Figure S2). Moreover, inhibition of bacterial growth was evident at 6h 30min of incubation, resulting in a 47.4% cellular density reduction of the Cp13 mutant grown in the LI condition compared to the HI condition (Fig. 3B). Similarly, we observed a 37% reduction of bacterial growth in the parental T1 strain following the addition of the iron chelator (Fig. 3A). In addition, an approximately 1-log reduction was observed in the CFU counts in the presence of the iron chelator when compared to the high iron media (Fig. 3C/D). As expected, the disruption of the *ciuA* gene in the *C. pseudotuberculosis* Cp13 strain resulted in a pronounced reduction in cellular density and growth rate relative to the parental T1 strain, indicating that the CiuA protein could be required for the sufficient acquisition of iron and normal growth of the mutant strain under iron limitation. These results are consistent with previous reports of growth analysis of *ciuA* mutants in *C. diphtheriae* strains, where bacterial strains that are defective in siderophore transport

often exhibit diminished growth under iron limitation and interruption of the *ciuA* gene greatly reduced the utilization of siderophore as an iron source [19].

DE analysis of iron-regulated genes

To estimate the relative gene expression across sample groups, pairwise analyses were conducted in counts of uniquely mapped reads of LI against HI cultures for each strain with the DESeq2 R package. Adjusted p-values were obtained using the Benjamin and Hochberg false discovery rate (FDR) method, and an adjusted p-value of 0.05 and a log2fold change of 0.5849 (fold change of 1.5) were set as cutoff for significant differential expression. A total of 106 (T1) and 80 (Cp13) genes were found to be differentially expressed between LI and HI conditions, of which, 35 genes were commonly expressed in the wild-type T1 strain and Cp13 mutant (Fig. 4A). In the wild-type T1 strain, 19 genes were identified up-regulated and 47 were down-regulated, while 41 genes were up-regulated, and 13 genes were down-regulated in the mutant strain (Fig. 4B). The distribution and expression profile for each strain were represented by a scatterplot of all DEGs expressed with a log2fold change of 0.5849/-0.5849 at a 5% FDR (Fig. 4C/D). Complete DE gene lists for each of the strain are provided in Additional file 3. To further analyze the identified DEGs, we performed GO and GO enrichment analyses. Functional annotation was retrieved from the GOfeat platform and enrichment analyses were conducted using the STRING tool with an FDR threshold set to 1%. See Additional file 5: Tables S9 and S10 for the full listing of GO terms identified for the DEGs of each strain. String was also used to construct functional interactions between predicted protein of the DEGs identified (Fig. 5).

Notably, a hierarchical clustering of the commonly expressed genes in both strains revealed a similar pattern of gene expression between wild-type and mutant strains, with 25 out of the 35 DEGs being consistently up- or down-regulated in the T1 strain and Cp13 mutant (Fig. 6). We believe that these DEGs may contribute directly to the homeostasis of iron in *C. pseudotuberculosis*. GO characterization of these overlapped DEGs revealed an overrepresentation of functional categories associated with the acquisition and transportation of hemin (*htaA*, *htaC*, *htaF*, *htaG*, *Cp_3070*, *Cp_3075*, *hmuT* and *hmuU* genes) and electron transfer activity of the oxidative phosphorylation process (*qcrCAB* and *ctaCEF* genes) (Fig. 6). In addition, PPI analysis also revealed significant functional annotation enrichments (FDR < 1%) for genes encoding subunits of the oxidative

phosphorylation pathway (KEEG 00190) and genes encoding proteins with a HtaA hemin-binding domain (PFAM PF04210) (Fig 5A/B). See Additional file 5: Figure S4 for protein enrichment analysis. These genes formed two significant and distinct protein clusters with a high PPI interaction value (over > 0.7), indicating significant biological association within the predicted products of these DEGs. Interestingly, two genes with DNA biding transcription factor activity, which are known to participate in corynebacteria iron homeostasis, were also up-regulated in both strains in response to iron restriction: the transcriptional repressor of iron protein A (*ripA*) and the DNA-binding regulatory protein HrrA (of the *hrrSA* two-component system) (Fig. 6). In *C. glutamicum* and *C. diphtheria*, these DNA binding regulators were shown to regulate iron and heme-containing proteins in response to iron availability [34,35].

Although a common expression pattern was observed between the strains, specific differences were also observed indicating a differential adaptive response between the mutant and parental strains under iron limitation. Among the unique DE genes identified, ribosome biogenesis (KEEG 03010, p adjusted value = 5.83e-07) and the tricarboxylic acid cycle (TCA cycle – KEGG 00020, p adjusted value = 0.000297) were the most affected cellular processes under iron restriction in the parental T1 strain (Fig. 5A). Overall, the T1 strain showed a greater number of DEGs down-regulated with overrepresentation of biological processes, including translation (GO:0006412) and TCA cycle (GO:0006099) (Additional file 4: Table S7). The Cp13 mutant exhibited an increased expression of DEGs in the biological process category assigned to the GO terms translation (GO:0006412) and regulation of transcription (GO:0006355), as well as over-represented genes assigned to the term DNA binding transcription factor activity (GO:0003700) (Additional file 4: Table S8). Ribosome biogenesis was the only category with significant enrichment among the unique DE genes in the Cp13 mutant (KEEG 03010, p adjusted value = 0.00289) (Fig. 5); however, contrary to the down-regulated ribosomal proteins in the parental strain, the genes assigned to this functional category were up-regulated in the mutant strain (Additional file 4: Table S8). A closer look at these unique DE genes reveals an interesting picture of the iron-adaptative response between the strains. Cultures of the Cp13 mutant strain under iron limitation produced a set of up-regulated transcripts involved in protein synthesis activity (*rpsO*, *rpsB*, *rpsP*, *rspT*), the transcriptional regulation of iron-sulfur clusters (*sufR*) and oxidative response (*WhiB4*), transient metabolic alteration favoring the generation of ATP through a micro-aerobic

pathway and rpf-induced growth factor (*rpfB*). In contrast, the majority of the unique down-regulated DEGs identified in the T1 strain were involved in pyruvate metabolism (*aceF*, *lpd*), TCA cycle (*sdhC*, *sdhB*, *sdhA*, *lpd* and *aceF*), ATP production (*atpF*, *atpH*), protein synthesis (*rplJ*, *rplL*, *rplM*, *rpmA*, *rpsC*, *rpsI*, *rpsL*, *rpsM*, *tsf*, *fusA*) and rpf-induced growth factor (*rpfA*) (See Additional file 3). In agreement with these results, the gene coding for the Cyclic AMP (cAMP) GlxR global transcriptional regulator was up-regulated in the T1 strain during iron restriction. In high G+C bacteria, GlxR is one of the most important transcription factors characterized as a dual transcriptional regulator controlling diverse cellular functions [26,27]

Iron limitation increased the transcription of putative hemin acquisition systems

Under iron limitation we observed an up-regulation of genes encoding surface exposed proteins with an enriched high-affinity HtaA hemin binding domain associated with hemin ABC-type transporters, suggesting that hemin may be a relevant iron source in C. pseudotuberculosis. Notably, all genes encoding proteins with an HtaA domain, *htaC*, *htaA*, *htaF*, *htaG* and *Cp_3070*, were up-regulated in both strains in response to iron limitation and were directly annotated to the term integral component of membrane (GO:0016021). The acquisition of iron bound especially to heme is very common in pathogenic bacteria and for many Gram-positive species, heme represents the preferred iron source, accounting for approximately 70-75% of the total available mammalian iron pool [34]. Overall sequence identity of the hemin-binding proteins is low among *Corynebacterium* species [35]; however, the proteins encoded by the *htaA*, *htaC*, *htaF*, *htaG* and *Cp_3070* genes were all identified as having at least one HtaA domain of approximately 150 amino acids with two conserved tyrosine residues, which are known to be critical hemin-associated iron ligands in *C. diphtheriae* [36] (Additional file 6: Figure S5). Alanine substitution of the tyrosine residues in the HtaA domain results in a significant reduction in hemin and hemoglobin binding by these proteins [35]. Interestingly, all hemin-binding proteins were found to be encoded by three distinct genetic loci with high-confidence protein interactions amongst the predicted products of these genes. These three loci comprise the well described hemin uptake system formed by the *hmuTUV* cluster associated with the *htaA* and *htaC* genes, and two genetic loci formed by the *htaF-htaG* and the *Cp_3075-Cp_3070* genes (Fig. 6). In *C. pseudotuberculosis*, the *htaA* gene appears to be organized in a single operon with the *hmuTUV* genes, while the *htaFG* and the *Cp_3075-Cp_3070* genes appear to be organized

in two distinct two-gene operons in the genome. All three loci exhibited a remarkable difference in expression levels between the up-regulated genes in both C. pseudotuberculosis strains. However, the *htaFG* and the *Cp_3070-Cp_3075* genes were more highly expressed (over 2 times) in iron-limited cultures of the Cp13 strain than in iron-limited cultures of the parental T1 strain, suggesting that the disruption of the *ciuA* gene could have elicited compensatory mechanisms involving the acquisition of iron in the form of hemin in the mutant strain.

Remarkably, the predicted products of the *Cp3075-Cp3070* and *htaFG* loci share striking structural similarities to the hemin and hemoglobin binding proteins CirA-ChtC (DIP0523-DIP0522) and ChtA-ChtB (DIP1520-DIP1519) described in C. *diphtheriae* [36]. In both species, all four genes are arranged in two-gene operons and encode proteins with N-terminal secretion signals, C-terminal transmembrane regions and conserved HtaA domains critical for binding hemin (Additional file 6: Figure S6). Moreover, the Cp_3075 shares 32.7% amino acid sequence identity and 57.8% similarity with CirA, while the Cp3070 protein shares 33% of amino acid identity and 50.4% similarity with the ChtC protein. The *Cp_3070* and *Cp_3075* genes are separated by an intergenic region of 12 nucleotides, indicating that these genes are probably regulated by a single promoter. Similar to the CirA protein of *C. diphtheria*, the Cp_3075 protein lacks a putative HtaA domain, however its strong association with the HtaA-like membrane protein encoded by the *Cp_3070* gene (PPI confidence score of 0.829 in both strains) suggests that this protein is likely involved in the uptake of hemin in *C. pseudotuberculosis* (Fig. 5A/B). The HtaF and HtaG proteins of C. pseudotuberculosis share 60% and 62% of amino acid sequence identity, 74% and 75.4% sequence similarity with the ChtA and ChtB proteins of *C. diphtheria*, respectively. In this study, both loci were strongly induced under iron limitation, which is consistent with the iron-DtxR regulation observed in the promotor region of the *chtAB* and *cirA-chtC* operons of *C. diphtheriae* [36]. Additionally, a DtxR regulatory site was identified upstream of the start codon of the htaF gene that shared 14 conserved residues with the DtxR binding site sequence identified for the *chtA-chtB* operon, confirming that the *htaFG* cluster is iron and DtxR regulated (Fig. 7). Despite being strongly up-regulated in response to iron restriction, we found no significant evidence of a regulatory DtxR sequence upstream of the putative *Cp3075-Cp3070* operon. Nevertheless, it is important to point out that the absence of a regulatory site could be partially attributed to the site prediction methodology, which restricted the

detection to highly similar regulatory sites previously identified in orthologous target genes. This is in accordance with the site prediction for the *cirA-chtA* operon, which showed a relatively poor match to the consensus DtxR-binding site, in spite of showing a strong iron and DtxR transcriptional regulation of the cirA promotor region [36]. In addition, the *chtA-chtB* operon of *C. diphtheriae* was shown to be flanked by inverted repeats of identical insertion sequence (IS) elements, creating a transposon-associated system. Both the *Cp3075-Cp3070* and the *htaFG* clusters, were identified within genomic islands of *C. pseudotuberculosis*, indicating that both clusters were probably acquired by horizontal gene transfer (Additional file 6: Table S11). A closer examination revealed that these putative genomic islands harbor several genes whose products are known to contribute to the virulence of various pathogenic bacteria (Additional file 6: Table S11). Although the contribution of these hemin-binding proteins to the virulence or survival of *C. pseudotuberculosis* remains to be determined, the relation between iron acquisition and virulence was previously reported by Billington et al. (2002) [20], which identified the *fagABC* iron uptake operon located in a pathogenicity island along with the primary virulence factor phospholipase D (Pld). The *C. pseudotuberculosis fagB(C)* mutants showed reduced abscess formation in CLA models, with the in vivo expression of the *fagABCD* genes significantly contributing to the virulence of the strain. The hemin uptake ChtAB system could also be implicated in the virulence of *C. diphtheriae*, being found in related strains responsible for the diphtheria outbreak in the 1990s [36,37].

Similarly, in both strains under iron limitation, we observed an up-regulation of the genes encoding a complete hemin iron acquisition system formed by the *htaABC* and *hmuTUV* genes [38]. In *C. pseudotuberculosis*, the *htaA* and *hmuTUV* genes appear to be organized in a single operon, while the *htaB* and *htaC* genes are probably regulated by two distinct promoters. In *C. diphtheriae*, the HtaC protein homologs mainly remain in the bacterial cell, while the HtaA protein was shown to be secreted in low iron cultures [38]. Three distinct putative 19-bp DtxR binding sites were identified upstream of the *htaAhmuTUV* operon and upstream the start codon of the *htaB* and *htaC* genes (Fig. 7), confirming the iron-DtxR regulation for the cluster [38]. Notably, the *hmuT* (hemin binding protein) and *hmuU* (permease) genes of the hemin utilization (*hmu*) operon were among the up-regulated genes in response to low iron availability in both strain (Table 1). The cluster of six genes is likely involved in the acquisition and transportation of hemin, sharing substantial sequence similarity with genes encoding hemin uptake systems in other

corynebacteria, including *C. diphtheriae* and *C. ulcerans*. An alignment of the *C. pseudotuberculosis* HmuT and HmuU amino acid sequences demonstrates a high level of identity with the Hmu components of *C. ulcerans* and *C. diphtheriae*. We observed a sequence identity of 97% and 98% with HmuT and HmuU of C. ulcerans and 81% and 82% of identity with HmuT and HmuU of. *C diphtheriae* (Additional file 6: Figure S7/S8). Moreover, the Tyr233 and Tyr347 conserved tyrosine residues were identified in the HmuT amino acid sequence (Additional file 6: Figure S8). These conserved tyrosine residues have been well characterized in many heme proteins and were shown to strongly bind and stabilize hemin Fe3+ [39–42]. In C. diphtheriae, it has been proposed that hemin is acquired from hemoglobin by the surface-exposed hemin binding HtaA protein, which sequentially transfers the heme group to the membrane HtaB protein and the substrate binding HmuT protein. The lipoprotein HmuT takes the hemin to the ABC transporters, which utilizes energy from ATP hydrolysis to pump hemin through permeases located in the cellular membrane [43,44]. As observed in other *Corynebacterium* species, our data indicate that iron and DtxR regulates the expression of the htaA and the htaC genes associated with hmu cluster (genes *hmuT* and *hmuU*), which are likely involved in the assimilation of hemin as an iron source. Consistently, a recent exoproteome study has identified significant levels of the HtaA, HmuT and HmuU proteins in the supernatant of *C. pseudotuberculosis* culture after experimental infections in the murine model, suggesting a link between hemin uptake proteins and virulence of this important pathogen [45].

Down-regulation of heme and iron-proteins under iron restriction

Collectively, iron restriction resulted in the down-regulation of genes encoding protein subunits of metabolic complexes involved in energy metabolism, with the down-regulation of 17 genes associated with the oxidative phosphorylation process (KEGG 00190) and the TCA cycle. The fact that the majority of these down-regulated gene products encode heme or iron-sulfur proteins suggests that intracellular restriction of iron usage is key to the iron starvation response of *C. pseudotuberculosis*. Oxidative phosphorylation is the final stage of energy-yielding metabolism in aerobic respiration and represented the most significant enrichment observed across all down-regulated genes identified in both strains (FDR < 1%), suggesting that iron limitation hinders *C. pseudotuberculosis* aerobic respiration. In both strains, we identified two gene clusters of the cytochrome *bcc-aa3* super complex, Ubiquinol-cytochrome C reductase – *qcrCAB*

operon and cytochrome oxidase c – *ctaCDEF* genes, with down-regulated genes subunits *qcrA*, *qcrB*, *qcrC*, *ctaC*, *ctaE* and *ctaF* (Table 1). The predicted products of these genes were all assigned to the GO term integral component of membrane, heme binding, iron ion binding and cytochrome-c oxidase activity. The *ctaE* gene is located adjacent of the *qcrCAB* genes clustered in the order *ctaE-qcrCAB*, suggesting that this gene could be transcriptionally regulated with the *qcrCAB* operon. In actinobacteria, the *ctaE* gene encodes subunit III of the cytochrome *aa3* oxidase, the *qcrCAB* genes encode the three subunits of the cytochrome bcc complex and the *ctaCDF* genes encode subunits (II, I and IV) of the cytochrome *aa3* oxidase complex (complex IV), forming a respiratory super complex chain in many high-GC Gram-positive bacteria [46,47]. The proteins in the *bcc-aa3* complex contain four redox prosthetic groups, formed by two heme b groups (QcrB), two c-type heme (QcrC), a 2Fe-2S cluster (QcrA) and two heme a groups (CtaD) in the *aa3* oxidase molecules [46,48]. In addition, we also observed the down-regulation of iron-sulfur protein genes of complex II of the respiratory chain amongst the DEGs of the T1 strain. The *sdhA*, s*dhB* and *sdhC* genes form the respiratory complex II of the respiratory chain in which succinate reduction acts as a pivotal link between the TCA and the oxidative phosphorylation process [46]. The *sdhCAB* gene cluster is formed by a membrane protein (SdhC) that interacts with ubiquinone and anchors the catalytic domain at the membrane surface. The flavoprotein (SdhA) and iron-sulfur cluster (SdhB) subunits catalyze succinate reduction [49]. The *sdhB* encodes a protein with three prosthetic Fe-Su clusters [48,50]. Taken together, these genes comprise the majority of down-regulated genes encoding proteins that require iron or heme as a cofactor, corroborating previous observations that there is a decrease in activity of proteins and protein systems which require heme or iron for their activity under iron limiting conditions (Fig. 5) [51,52]. This is further confirmed by the down-regulation of the *ftn* gene in the T1 strain, which encodes a ubiquitous iron storage protein. In *C. pseudotuberculosis*, the *ftn* gene encodes a spherical protein (Ferritin) of 170 amino acids with a large hollow center that acts as iron-storage cavity. Previous results have shown that under iron restriction, *ftn* expression is reduced in order to control the intracellular iron demand of the bacterium [15].

Iron restriction hinders energy metabolism

Generally, low levels of iron are expected to activate a cellular iron restriction response involving the down-regulation of genes encoding key iron/heme proteins of metabolic processes, resulting in a state of low metabolic activity in which bacteria can adapt to the

hostile iron-poor environment [15]. Accordingly, most strain-specific genes down-regulated in the T1 strain under iron-restriction conditions encoded components of the oxidative phosphorylation process as well as enzymes of the TCA cycle. These down-regulated genes included succinate dehydrogenase iron-sulfur proteins (genes *sdhA, sdhB* and *sdhC*), dihydrolipoamide acyltransferase (*aceF*) and dihydrolipoamide dehydrogenase (*ldh*). Concomitantly, expression of the genes encoding subunits B (*atpF*) and δ (*atpH*) of the *F1F0*-complex was also decreased in the T1 strain, indicating reduced ATP synthesis associated with the respiratory pathway. The *F1F0*-complex integrates the electrochemical gradient of proton translocation through the membrane for the synthesis of ATP [46]. ATP synthase is encoded by the *atp* operon, which is tightly regulated by cellular ATP requirements and respiratory chain activity. Regulation of ATP synthase is critical for the survival of *M. tuberculosis* inside the macrophages, which has been shown to be down-regulated upon infection in order to lower the metabolic rate of the pathogen in a nutrient-poor environment [53].

In contrast, we have observed an up-regulation of the ndh gene in the Cp13 mutant, which encodes an iron-free type II NADH dehydrogenase. The copper containing NADHII and the cytochrome *bd* oxidase complex are responsible for the generation of a less efficient proton gradient during micro-aerobic growth, which in turn generates ATP through ATP synthase [54,55]. Although a shift of electron flow from the *bcc-aa3* complex to the *bd* might lead to a decrease rate of ATP synthesis and metabolic activity, these results thus suggest that, in contrast to the parental T1 strain, the disruption of the *ciuA* gene in the Cp13 mutant favored the use of iron-free proteins of the micro-aerobic chain under iron restriction. This is in accordance with previous reports that have shown that this respiratory micro-aerobic chain is predominant when iron is scarce [56]. Nonetheless, this shift in metabolic activity in the Cp13 mutant could promote ATP production under iron limitation, providing a mechanism for the bacterium to compensate for the restriction of intracellular iron demand with the ATP required for pivotal biosynthetic processes, such as protein synthesis. Since translation is an energy-consuming process, there is a strong connection between ATP availability and protein synthesis [57], which is in accordance with the down-regulation of several ribosomal protein genes exclusively in the Cp13 mutant.

Proteins synthesis under iron limitation

Under iron restriction, the strains responded differently regarding the expression of genes encoding ribosomal constituents, with all ribosomal DE protein genes being repressed in the T1 strain and up-regulated in the Cp13 mutant. In the T1 strain, the *rplJ*, *rplL*, *rplM*, *rpmA*, *rpsC*, *rpsI*, *rpsL*, *rpsM* genes were down-regulated by over 1.5-fold difference together with the other down-regulated genes *fusA* and *tsf*, which encode the elongation factors EF-G and EF-Ts associated with translation. In the Cp13 mutant, the genes encoding ribosomal proteins of the small 30S ribosome subunit (genes *rpsB*, *rpsO*, *rpsP*, *rpsT*) were identified as up-regulated by >1.5-fold difference (See Additional file 3). Under most growth conditions, the rate of protein synthesis is strongly associated with efficient ribosome biogenesis, which is in turn tightly coupled with the energy cell production and nutrient availability [57]. However, in the mutant strain, we observed up-regulation of a set of genes required for protein synthesis whose expression is normally down-regulated when under nutritional restriction [57]. Although unexpected, the increased expression of ribosomal-related genes might reflect an attempt to continue protein synthesis, which was correlated with the higher fold-expression of genes encoding hemin-binding proteins and the metabolic activity of the Cp13 mutant strain compared to T1 (See Additional file 3).

The expression of transcription factors under iron restriction

The transcriptomic analyses showed the upregulation of two well described iron-dependent TFs amongst the differentially up-regulated genes in both strains: the regulator HrrA of the two-component system (TCS) hrrSA and the RipA repressor (Fig. 6), which are known to be regulated by the availability of iron in other corynebacteria [58,59]. The hrrA gene was up-regulated by a 1.6-fold and 2-fold difference in the T1 and Cp13 strains, respectively (Table 1). The two-component system (TCS) are ubiquitous bacterial systems composed of two proteins: a membrane-associated sensor histidine kinase (HK) and a response regulator (RR), which detect specific environmental signals allowing bacteria to regulate osmolarity, nutrient acquisition, modulate gene expression and modify protein interaction [60]. *C. pseudotuberculosis* encodes 8 two-component systems, but little is known about their function. The DNA-binding response regulator HrrA of the hrrSA system in *C. pseudotuberculosis* shares high amino acid sequence identity with the homologs HrrA regulator of *C. diphtheriae* (91%) and of *C. glutamicum* (83%) (Additional file 6: Figure S9). In *C. diphtheriae* and *C. glutamicum*, the *hrrSA* system is responsible for the activation of the heme oxygenase (*hmuO*) promoter and for

the repression of genes involved in heme homeostasis [43,61,62]. The *hrrSA* system has also been shown to regulate the expression of the *ctaE-qcrCAB* and *ctaCF* genetic clusters [61]. In *C. glutamicum*, a *ΔhrrA* mutant showed a strong decrease in the expression of the *qcrCAB* and *ctaCEF* genes in the presence of heme, which suggests an *hrrSA* regulation. In addition, genes belonging to both clusters were also shown to be down-regulated in the presence of iron in *ΔhrrA* mutant [61].

Correspondingly, the *ripA* gene was up-regulated by a 2-fold difference in the T1 strain and by a 2.4-fold in the Cp13 mutant strain (Table 1). It has been demonstrated that under iron limitation, the RipA protein reduces cellular iron demand by acting as a repressor of several genes encoding iron-containing proteins, such as aconitase and succinate dehydrogenase [59,63]. These results are in accordance with the down-regulation of the iron-sulfur proteins encoded by the *sdhCAB* genes in the T1 strain (Additional file 3: Table S5).

Consistent with a lower metabolic activity, the global regulator GlxR was up-regulated by 1.6-fold in the T1 strain under iron restriction. In vitro binding assays and in silico work indicate that GlxR could be involved in the direct control of over 200 genes in *C. glutamicum* [26] and 80 genes in *C. diphtheriae* [56]. In these species GlxR regulates the expression of genes involved in carbohydrate metabolism, aerobic and anaerobic respiration, fatty acid biosynthesis, deoxyribonucleotide biosynthesis, cellular stress response and resuscitation [27]. In *M. tuberculosis*, a GlxR homolog was reported linked to bacterial survival inside macrophages [57]. Other induced regulators identified and related to the molecular function term DNA transcriptional factor and DNA binding among the up-regulated genes in the Cp13 mutant were the ArsR family repressor of iron-sulfur cluster biogenesis gene (*sufR*) and the WhiB family transcription regulator (See Additional file 3). The up-regulated *whiB4* gene encodes a protein with a redox-sensitive iron-sulfur cluster that is potentially associated with the regulation of genes involved with the oxidative stress response of *C. glutamicum* and *M. tuberculosis* [64,65]. This iron-sulfur containing transcription factor was also implicated in the intracellular replication of *Mycobacterium marinum* in macrophages and diminished virulence [66]. Furthermore, overexpression of the *whiB4* gene (*whcA* homologous) is strongly associated with slow growth in *C. glutamicum* cells [64].

Regulation of resuscitation promoting factors under iron limitation

Under iron limitation, we observed an up-regulation of the resuscitation-promoting factor B *(rpfB*) gene in the Cp13 mutant and the down-regulation of the resuscitation-promoting factor A (*rpfA*) gene in the parental T1 strain (See Additional file 3). These Rpf proteins contain a highly conserved Rpf domain with peptidoglycan-hydrolyzing activity [67,68]. In *M. tuberculosis*, the Rpf proteins act as bacterial growth factors that stimulate the growth of dormant mycobacteria cells, contributing to *M. tuberculosis* persistence and virulence inside the host [69]. The *rpf* genes are also involved in bacterial growth and culturability of nonsporulating bacteria [69]. The role of these independently regulated genes is still inconclusive with no significant phenotype being assigned to a specific *rpf* gene, suggesting a functional redundancy between these genes [70]. How exactly the iron acquisition *ciu* cluster and the *rpf* genes are associated is not immediately clear; however, the up-regulation of the *rpfB* gene in the Cp13 mutant indicates an attempt to grow in a media unconducive for optimal replication.

Regulation of siderophore systems under iron limitation

We selected the Cp13 mutant as an iron limitation model hoping to illuminate how this bacterium copes with iron stress, and to understand the role that the siderophore *ciu* cluster plays in iron acquisition. Unexpectedly, we were unable to address the involvement of siderophores in the acquisition of iron by *C. pseudotuberculosis*. Siderophore synthesis has a pivotal role in iron acquisition for many pathogenic species [71]. For these pathogens, previous results using *ciuA* mutants have highlighted the importance of this protein in virulence and pathogenesis [16,21,45]. Surprisingly, none of the genes believed to encode proteins involved in siderophores biogenesis and uptake (*ciuABCDEF* and *fagABC-fagD* clusters) were identified as differentially expressed in the parental T1 strain, or in the Cp13 mutant strain. Contradictorily, regulatory DtxR sites were identified upstream the genes *ciuD* and *ciuE* of the *ciuABCDEF* genetic cluster and upstream the *fagA* gene of the f*agABC* cluster (Fig. 7), indicating that these clusters are probably regulated by DtxR and iron. This could suggest that similar to other iron-regulated genes, these putative siderophores clusters must be actively repressed by DtxR under sufficient iron availability. However, under iron limitation, gene expression is likely to require additional regulatory pathways [43,72–74]. This hypothesis was first proposed by Billington et al. (2002) [20], where the expression of the siderophore ABC-type transporters encoded by the *fag* genes were only observed in vivo, thus suggesting

that unknown host factors could play an important role in the expression of genes involved in siderophore synthesis and uptake. Although no differential expression was observed, the *ciu* cluster appears to be essential for the growth and the bacterial adaptive responses under iron restriction. However, the data on siderophore production and acquisition in *C. pseudotuberculosis* is still circumstantial and further studies must be conducted in order to clarify the role of high-affinity siderophores in the acquisition of iron in *C. pseudotuberculosis*. Comparing the transcriptome profile of the strains, we were able to identify a shift in metabolic activity of the mutant strain favoring ATP production through iron-independent pathways. Since the majority of the ATP in the cell fuels protein synthesis [57], our results have demonstrated that the disruption of the *ciuA* gene induced an expression response involving the up-regulation of genes responsible for the uptake of iron in an attempt to compensate for the absence of a functional siderophore uptake system. Given the redundancy of iron utilization systems and the importance of iron acquisition, the disruption of the *ciuA* gene in the Cp13 mutant could be masked by a non-siderophore iron uptake system. This hypothesis is further supported by the remarkable 2-fold difference of hemin uptake systems between the Cp13 mutant strain and the parental T1 strain (Table 1).

Transcriptional regulatory network of the DEGs identified under iron restriction

The transcriptional regulatory network of bacteria is a complex biological system that involves the regulation of gene expression by the binding of specific transcription factors (TFs) to their DNA sites in response to external signals, allowing bacteria to rapidly adapt to changes in the environment. Since the majority of genes encoding DNA-binding transcription regulators are highly conserved in corynebacteria, we applied a transfer of conserved regulatory interactions of known transcription regulators from the taxonomically closely species *C. diphtheriae, C. glutamicum* and *M. tuberculosis* to the genome of *C. pseudotuberculosis*. Binding motifs of homologous TFs were used to create an HMM profile that was used to search for the presence of regulatory sites upstream of the coding regions of target homologous genes in the *C. pseudotuberculosis* genome. Although this conservative approach does not identify new regulon members, it is extremely reliable in eliminating a large number of false positives associated with in silico predictions. On the other hand, this approach is limited by the level of conservation between the references and the target species, where a regulatory interaction can only be transferred between the reference and the target genome if a conserved DNA binding site

is identified in the promoter region of an orthologous target gene (or operon) to which an orthologous TF has also been detected [29]. This approach yielded 187 interactions with 21 transcriptional factors and 119 transcriptional binding sites involved in the regulation of over 200 genes. Based on these results, we constructed interaction networks involving the differentially expressed genes regulated by iron restriction in both strains. Interaction network analysis using the DEGs generated 3 networks, with two networks formed by two or less than two DEGs and one network, related to energy metabolism and iron homoeostasis, involved in the regulation of 27 DE genes. Collectively, the iron network identified involved genes regulated by the global regulator GlxR protein and the master regulator of iron metabolism protein (DtxR - diphtheria toxin repressor) (Fig. 8). The majority of the differentially expressed genes identified in our analyses were shown to be regulated by DtxR (Fig. 8), which is not unexpected given its role in iron metabolism. DtxR is the key TF regulator of iron homeostasis in corynebacteria, while the iron-dependent repressor IdeR (DtxR homologous) regulates the iron homeostasis in *Mycobacterium* species. The DtxR metal-ion-dependent protein is composed of an N-terminal DNA binding domain and a flexible C-terminal domain with two metal-binding sites [76]. Although the breadth of the gene expression regulation may vary among the target genes in *Corynebacterium* species [58,77]; overall, in the presence of ferrous iron, DtxR mediates gene expression regulation by binding directly to a 19-bp operator sequence in iron-regulated promoters, thus preventing the binding of RNA polymerase. In contrast, in an iron deficient environment, DtxR is unable to bind Fe2+ and the repression of target regions is released [78]. Auto-regulation of global TFs is common among bacteria [59,79,80]. However, no regulatory site was located upstream of the DtxR gene, which indicates that the DtxR is not auto-regulated. In *C. pseudotuberculosis*, the homologous DtxR gene encodes a protein of 226 amino acids with over 79% of sequence similarity to the DtxR of *C. diphtheriae* and *C. glutamicum*, and 70% to the IdeR of M. tuberculosis. 20 putative DtxR binding sites with the consensus sequence TTAGGTTAGGCTAACCTAN were identified in the genome of *C. pseudotuberculosis* (Fig. 7), which might be involved in the regulation of over 40 genes encoding proteins involved in iron acquisition, storage, metabolism and transcriptional regulation.

Confirming previous findings, putative regulatory DtxR sites were identified upstream of the coding regions of genes related to iron acquisition and transportation (*htaAhmuTUV*, *htaC* and *htaF*); iron storage (*ftn*) and the ferric citrate transport system (*fecB*). Although

gene clusters involved in siderophore acquisition were not identified as differentially expressed under iron limitation, regulatory site motifs were found in the intergenic region between the *fagD* gene and the *fagABC* cluster, upstream the *fecCDE(CD)* operon and upstream the coding regions of the *ciuE* and *ciuD* (of the *ciuABCDEF* cluster) genes (Fig. 7) [19,81]. The proposed DtxR network revealed a complex hierarchical cross-regulatory structure, with DtxR negatively controlling three other transcription regulators identified as differentially expressed in our results: the *ripA*, *sufR* and *hrrA* (Fig 15). DtxR regulatory sites were identified upstream the coding regions of DNA-binding regulator *ripA* and *sufR*, confirming that both regulators are actively repressed by iron and DtxR. The *sufR* gene is a metalloregulator of the ArsR protein family responsible for the repression of the suf gene cluster, which is involved in the biogenesis of iron-sulfur clusters in bacteria [82]. The RipA protein is part of the AraC family of transcription regulators and it has been demonstrated that under iron limitation, the DtxR repression of the *ripA* gene is relieved and the transcription of RipA protein reduces cellular iron demand by acting as a repressor of several genes encoding iron-containing proteins, such as aconitase and succinate dehydrogenase [59,63]. Although no homologous regulatory *ripA* site was found upstream the *sdhCAB* genes, previous results have shown that the operon is actively repressed by the binding of the *ripA* repressor [59]. In both *C. pseudotuberculosis* strains, regulatory biding sites for DtxR and GlxR were identified in the promoter region of the *sdhCAB* operon (Fig. 8), indicating that the *sdh* operon is positively co-regulated by DtxR and negatively regulated by RipA and GlxR. In *C. glutamicum*, the iron-DtxR repression was also shown to regulate the expression of the *hrrA* promoter in an iron-dependent manner [61,62]. In contrast, no DtxR binding motif was reported in the promoter region of the *hrrSA* system in *C. diphtheriae*, exposing differences in regulation between *Corynebacterium* species [43]. Comparatively, we also found no significant evidence for a DtxR binding site in the promoter region of the *hrrA* gene. However, given the predicted divergence in DtxR binding sequence within the Corynebacterium species, this observation was probably due to the low conservation of the DtxR regulatory site among these two species. To confirm this hypothesis, we aligned the regulatory site of *C. glutamicum* with the in silico predicted site identified in *C. pseudotuberculosis* FRC41 strain by Trost et al. (2010) [58]. As expected, these two species shared 8 out of the 19-bp DtxR binding site predicted for the target *hrrA* gene. Furthermore, we conducted a BLAST search of the Cp13 and T1 strains genomes for the *in silico* predicted 19-bp DtxR binding site. The DtxR predicted site

(TCAAGTGACTCTCACGTAT) was located upstream of the start codon of *hrrA* gene overlapping the coding region of the *hrrS* gene, sharing 19-bp (out of 19) consensus identity with the predicted DtxR binding motif (Fig. 7). Besides the nucleotide diversity, these regulatory DtxR sites were both identified overlapping the *hrrS* gene, which could indicate that only the *hrrA* gene of the *hrrSA* system is regulated by iron and DtxR *in C. glutamicum* and *C. pseudotuberculosis* [61]. In *C. diphtheriae* and *C. glutamicum*, the *hrrA* gene of the *hrrSA* system is responsible for the activation of the heme oxygenase (*hmuO*) promoter and for the repression of genes involved in heme homeostasis [43,61,62]. In addition, *hrrA* mutants of *C. glutamicum* showed a strong decrease in the expression of the *qcrCAB* and *ctaCFE* genes, which could suggest that these genes might be regulated by the *hrrSA* system. However, genes belonging to both clusters were also shown to be down-regulated in the presence of iron, indicating that these clusters could also be regulated by iron [61].

The regulatory motif search for GlxR-binding sites in the genome of *C. pseudotuberculosis* yielded 18 potential transcriptional GlxR-regulated sites controlling directly 12 DEGs involved in aerobic respiration, stress response and resuscitation. The GlxR protein acts as a dual transcription regulator, positively controlling the expression of the *ctaEqcrCAB*, *ctaCF*, *ctaD* and *rpfB* genes, while represses the expression of the *sdh* operon (Fig. 8). A regulatory GlxR binding site was also identified upstream the transcriptional regulator RamB, indicating that GlxR could be cross-regulating the expression of the *fkpb* gene under iron restriction (Fig. 8). Interesting, we have also identified a coregulatory interaction between the RamB and RamA transcription factors and the *Cp_9550* gene (Fig. 8). In *C. glutamicum*, both RamA and RamB act controlling the utilization of different carbon sources and metabolic adaptation to specific nutritional conditions [75].

5.1.4. Conclusion

In summary, we identified a total of 106 and 80 transcripts differentially expressed between low iron and high iron cultures of the wild-type parental T1 strain and the Cp13 mutant of *C. pseudotuberculosis*, respectively. Of these total, 25 genes were found consistently up- or down-regulated in the parental and mutant strain under iron restriction and could potentially play a role in the intracellular iron homeostasis of *C. pseudotuberculosis*. Functional analysis of these genes demonstrated that the up-

regulation of systems involved in the acquisition of iron from hemin and the down-regulation of intracellular iron demand of iron-dependent aerobic respiratory pathways are key to the response of *C. pseudotuberculosis* to iron restriction. Notably, these hemin uptake systems were substantially expressed in the Cp13 mutant, highlighting the importance of hemin and suggesting that *C. pseudotuberculosis* does not restrict iron uptake to a single strategy (for example siderophores), but can adjust iron acquisition accordingly. Some of these hemin uptake genes and many known virulence factors were also identified harbored by genomic islands, corroborating with the association of iron uptake and virulence of pathogenic bacteria. Additionally, two transcription factors (*ripA* and *hrrA*) known to control the expression of genes encoding iron-containing proteins were also up-regulated in both strains. Comparative analysis of the transcriptome profile for the T1 wild-type and Cp13 mutant highlighted the differences between the iron-mediated responses of the two strains involving genes that were directly related to bacterial growth and metabolism. While we were unable to address the expression of siderophore-based iron acquisition systems for this bacterium, our results, although admittedly preliminary regarding the *ciuA* gene, indicate that siderophore-mediated iron acquisition may be required for optimal growth and could be involved in the adaptative gene expression response of *C. pseudotuberculosis* to iron restriction. To our knowledge, this is the first study to analyze changes in gene expression between high iron and low iron cultures of two strains of *C. pseudotuberculosis*, providing a comprehensive transcriptional analysis of this important pathogen. Our findings contributed to an enhance understanding of the molecular mechanisms involved in the *C. pseudotuberculosis's* ability to rapidly adapt to the restriction of iron, revealing potential gene targets that can be used in the development of effective therapeutic strategies for CLA.

5.1.5. Methods

Bacterial strains

The biovar ovis pathogenic wild-type T1 and the ciuA iron-acquisition-deficient Cp13 mutant were both obtained from the Laboratory of Cellular and Molecular Genetics (LGCM) at the Federal University of Minas Gerais/Brazil. The T1 strain was isolated from a caseous granuloma found in CLA-affected goats in Brazil (Bahia state) [16]. The Cp13 mutant was generated by Dorella et al. (2006) [16], using the in vivo insertional mutagenesis of the reporter transposon-based system TnFuZ in the strain T1. This system

combines a transposable element (Tn4001) with a modified Gram-positive bacterium alkaline phosphatase gene (*phoZ*), and it is used for the discovery and mutagenesis of cytoplasm exported proteins. The PhoZ reporter protein is activated upon its exportation from the cytoplasm and insertional mutant colonies are identified by the colorimetric reaction of the alkaline phosphatase with the 5-bromo-4-chloro-3-indolylphosphate substrate [17]. The molecular characterization of the Cp13 mutant showed that the insertion disrupted the ciuA gene, which encodes a putative-iron transport binding protein of the *ciuABCD* operon [16].

Media and growth analysis of *C. pseudotuberculosis* strains

Wild-type T1 and Cp13 mutant strains were grown individually either in the presence of the iron chelator 2,2'-dipyridyl-DIP (Low Iron condition), or without it (High Iron condition). Precultures and main (inocula) cultures for both conditions (LI and HI) were all prepared with Brain Heart Infusion (BHI) broth-0.05% (v/v) Tween® 80 (Sigma-Aldrich) under routine conditions. The iron-chelated BHI medium was prepared with 250 μM of ferrous iron chelator 2,2'-dipyridyl (Sigma Aldrich), which, due to its low aqueous solubility, was prepared with 40% (v/v) of ethanol (0.5M 2,2'-dipyridyl stock solution) and then was kept refrigerated and protected from light until its use. The 0.5M 2,2'-dipyridyl stock solution was diluted in sterile water to a concentration of 0.1M 2,2'-dipyridyl 8% (v/v). To chelate the iron content of the medium, 0.1M 2,2'-dipyridyl solution was added to fresh BHI broth to a final concentration of 250 μM 2,2'-dipyridyl with 0.02% (v/v) of ethanol. To normalize the effect on gene expression caused by the presence of ethanol, HI cultures were prepared with 0.02% (v/v) sterile-filtered ethanol (final concentration).

Bacterial cultures on iron-chelated (low iron) media and on non-chelated (high iron) media for the wild-type T1 and the Cp13 mutant strains were conducted as follows: precultures were prepared inoculating a single colony of the strains T1 and Cp13 harvested from fresh BHI agar plates. Precultures were incubated overnight at 37˚ C (140rpm). After incubation, precultures were diluted to a start optical density of 0.02 (OD600) and inoculated in BHI cultures prepared in the absence (HI) and in the presence (LI) of 250 μM 2,2'-dipyridyl. Cultures were incubated at 37˚ C (140rpm) and bacterial growth was monitored hourly by optical density (OD600). The iron-stressed condition was determined by the addition of the iron-chelator, and the stress was confirmed by a

reduction of bacterial culture density in the LI medium in comparison to bacteria grown in the HI medium after 6h and 30 minutes of incubation. Bacterial viability was assessed by the number of colony-forming units per milliliter (CFU/mL) after the 6h30min incubation. OD data were used to estimate cellular density by applying the final and initial OD-values (at t1=0 and t2=6h30min) to the equation: Growth rate (Gr)= log𝑓0 〖OD〗 _2-log𝑓0 〖OD〗 _1/(t_2- t_1). Growth rate and CFU counts represent averages of three independent assays.

In addition, to assess any possible toxic effect of ethanol under the described culture conditions, bacterial growth and viability were monitored hourly over a 14-hour period by optical density (OD600) and by the number of CFU/mL after 11 hours of incubation using the Cp13 mutant. Comparisons were made by inoculating a starting bacterial preculture at concentration of 0.02 (OD600) in BHI media prepared with 0.02% and 0.1% (v/v) of ethanol against a BHI ethanol-free inoculum.

Statistical analyses

Statistical analyses for the growth assays and bacterial viability under iron limitation and in the presence of ethanol were carried out using the GraphPad Prism v.5.0 software (GraphPad, SanDiego, CA, USA).

Bacterial transcriptome assays and RNA extraction

In order to identify the transcriptional response of **C. pseudotuberculosis** involved in bacterial persistence under the growth conditions described above, wild-type T1 and Cp13 mutant strains were grown in an iron-chelated media (low iron - LI) and in a non-chelated iron media (high iron - HI). Incubation of individual HI and LI samples were conducted in parallel and independent sets of paired samples (HI and LI conditions) were prepared for each strain. Bacterial growth was monitored hourly by optical density (OD600) and total RNA extractions were carried out following 6h30min of incubation. Following the determinate incubation period, 20 mL of each culture (LI and HI) was centrifuged for 10 minutes at 5000 rpm (4˚ C) in 50 mL conical tubes, and the supernatants were discarded. The bacterial pellets were kept on ice and resuspended in 300 µL of Tris/EDTA/SDS (Tris HCl pH8 50 mM/L–EDTA pH8 10 mM/L) lysis buffer per 10 mL of used initial culture volume. Cell suspensions were homogenized with 0.1%

of SDS and 1 mL of TRIzol® Reagent (Invitrogen™) was added (1 volume per 10 mL of used initial culture). Pellets from each tube were divided into 2 new tubes filled with 1mm diameter glass microbeads (Bertin Technologies). The cells were mechanically lysed using Prescellys®24, set for 2 cycles (15 seconds per cycle) at 6200 rpm with 5 seconds interval between each cycle. The tubes were immediately placed on ice following the lysis. The cooled samples were centrifuged for 10 seconds at 14,000 rpm and the supernatants were carefully transferred to a new tube for total RNA extraction. Subsequent extraction steps were carried out following TRIzol® Reagent manufacturer's recommended protocol and total RNA was eluted in 50 µl of RNase free water. Traces of contaminating DNA were eliminated with the Ambion® TURBO DNA-free™ DNase kit (Invitrogen), following manufacturer's protocol. DNase treated samples were eluted in 50 µl of RNase-free water and RNA quantification was done using NanoDrop NDZ1000 (Thermo Scientific, Wilmington, DE) spectrophotometer. Samples with a 260/280 ratio over 1.9 were considered pure and used in subsequent steps. Samples with a high concentration of total RNA were obtained using a vacuum concentrator and by combining the previously divided samples. RNA quality and integrity of combined samples were verified by the electrophoretic profile of the 16S and 23S rRNA subunits using Agilent™ 2100 Bioanalyzer™ instrument. A total of 14 samples (7 sets of paired samples) were selected for library construction: 3 paired sets (HI and LI condition) of RNA samples from the wild-type parental T1 strain and 4 paired sets (HI and LI) of RNA samples from the mutant Cp13. All RNA samples were subjected to ribosomal RNA depletion and purification using the RiboMinus™ Bacteria Transcriptome Isolation Kit with the RiboMinus™ followed by use of the Concentration Module (Invitrogen), as described in the manual.

Library preparation and sequencing

Approximately 5 µg of purified depleted RNA samples were used to prepare 14 individual single-end libraries with the Ion Total RNA-seq v2 for Whole Transcriptome Library kit (Thermo Scientific). Purified RNA was enzymatic fragmented using RNase III. The fragmentation procedure produced a distribution of RNA fragment sizes that ranged from 35 to few thousand nucleotides, with an average size of 100-200 nucleotides. The size distribution and yield of cDNA were assed using the Agilent™ 2100 Bioanalyzer instrument and NanoDrop™ spectrophotometer. Template preparation was carried out with the amplified cDNA and each library template was clonally amplified on the Ion

Sphere™ Particles using the Ion OneTouch™ 2 System. Two rounds of sequencing were performed on the Ion Proton™ platform. The library preparation and sequencing steps were performed at the Federal University of Parana/Brazil.

Quality assessment and differential expression analysis

The output raw data files generated by the Ion proton system were converted to FASTQ-sanger file format using the Torrent Suite Software v. 5.0.5 (Thermo Scientific). The software was used to remove low quality bases at the 3'-end and the ion adaptor sequences. Raw data was then quality checked using the FASTQC tool (Babraham Bioinformatics, Cambridge, UK) [22]. Per base quality dropped significantly at the 3'-end extremity and a sliding window trimming was performed by Trimmomatic; cutting once the average base quality fell within the Phred score of 10 (90% accuracy) [23]. Trimmed reads were again quality checked using the FASTQC tool and mapped to the Corynebacterium pseudotuberculosis strain 1002B genome (NZ_CP012837.1), using the Torrent Mapping Alignment Program v3.4.1 (TMAP) [24]. Mapping parameters were set as default and the -mapall command was used, combining all 4 available mapping algorithms (map1, map2, map3, map4) to successfully address the variable read length and the high homopolymer error rate associated with the Ion Torrent technology [25]. The abundance of unambiguously mapped reads to a gene feature in the aligned genome was determined using HTSeq v0.8.0 package [26]. Differential expression analysis was performed with the bioconductor R package DESeq2 v.1.16.1 [27]. In short, using a generalized linear model (GLM) for each gene, the uniquely mapped read counts were modeled as following a negative binomial distribution with mean and dispersion, where the mean is proportional to the concentration of RNA fragments counted, scaled by a size factor normalization, which accounts for differences in sequencing depth between samples [27]. The linear model fit generated coefficients that indicate the overall gene expression strength and the log2-fold change between the low iron condition (LI samples) compared to the reference (HI samples). Pairwise comparisons were conducted in the normalized count data and significance was determined by the Walt test with p-value adjusted for multiple testing using the Benjamini and Hochber procedure [27]. Genes with a p adjusted value < 0.05 (FDR 5%) were considered to be differentially expressed (DE); these DE genes were ranked by their fold-change and were considered as being up- or down-regulated when there was a minimum 1.5-fold change difference with respect to the HI samples. Regularize-logarithm transformed (rlog) data was used to present relative

gene abundance using the rlog function and applied to measure sample to sample distance by calculating the Euclidean distance between samples and by Principal Component Analysis (PCA). Visual representation was done using the pheatmap v.1.0.8 and ggplot2 v.2.2.1. A heatmap of the differentially expressed genes identified in both strains was hierarchically clustered using the rlog of the normalized count data. A volcano plot was designed for both strains, demonstrating the fold change (log2 ratio) against the statistical significance (-log10 adjusted p-value) of the counted genes. A Venn diagram of differentially expressed genes was created to show the numbers of common and distinct genes identified in both strains. Batch effect was addressed in the sequenced data samples of the Cp13 mutant after PCA evaluation and it was attributed to sample processing. Strain-specific data samples were processed in separate groups and batch identification was included into the DESeq2 design formula [27].

Functional annotation and PPI networks of differentially expressed genes (DEGs)

The biological function of differentially expressed genes was retrieved from the GOfeat platform. In addition, protein-protein interactions (PPIs) analyses were conducted to investigate direct and functional associations between differentially expressed gene products. Protein interactions were constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) v.10.5 and were based on the transfer of functional annotation between orthologous genes [28]. Protein interactions with a combined score of >0.4 and a PPI enrichment value < 0.01 were considered significant. Enrichment analyses were also carried out using the STRING tool and were set to a 1% false discovery rate (FDR).

DEGs regulatory networks and genomic islands (GIs) prediction The C. pseudotuberculosis transcriptional regulatory network (TRN) was predicted using regulatory interactions of experimentally confirmed networks identified in the taxonomically related species: *C. glutamicum* (NC_003450) *C. diphtheriae* (NZ_LN831026) and *M. tuberculosis* (NC_000962). This methodology combines the detection of orthologous transcription factors and target genes with the identification of conserved binding motifs to reliably transfer transcriptional regulatory interactions between related species [29]. First, orthology detection was performed using bi-directional BLAST with e-value cutoff of 10-35, which was previously stablished to adequately cluster homologous proteins within the actinobacteria phylum. [30]. Once the

homologous target genes were identified, we scanned the promoter region of these genes (or operons) (+20 to -560 relatives to the start of the coding region of the gene) for conservative binding regulatory sites. Experimentally confirmed regulatory sites were used to detect in silico conserved binding motifs in the genome of *C. pseudotuberculosis* using Hidden Markov Model (HMM) profiles, with HMMER software. The HMMER bit score was used to infer the confidence level (high/medium/low) of each predicted binding site, reflecting how well the site sequence matched the profile model [31]. The bit score is calculated as the log-odds ratio score (base two) comparing the likelihood that the query sequence is a significant match to the profile HMM to the likelihood generated by a random model. Predicted regulatory site motifs with a bit score equal or greater than the selected score threshold of 3.45 were considered true positives and indicate high confidence site prediction. Predictions with a bit score between 0 and <3.45 were considered medium, while negative scores mean that the sequence is probably similar by chance (rated low) [32]. Networks were constructed using the DEGs identified in our transcriptome analysis.

Gipsy (Genomic Island prediction software) [33] was used to predict genomic islands (GIs) in the genome of the parental wild-type T1 strain of *C. pseudotuberculosis* against the closely related nonpathogenic genome of *C. glutamicum* (NC_003450). Comparatively, island prediction was conducted in the reference genome of the *C. pseudotuberculosi*s 1002B strain (NZ_CP012837.1). We looked at the genes in the putative predicted islands to see if they were part of the differentially expressed genes that we identified earlier in the analysis.

Availability of supporting data

The datasets generated and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository under the accession number GSE114125 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114125>. The datasets supporting the conclusions of this article are included in this article and its supplementary information files.

5.1.6. Additional files

Additional file 1: Table S1. Sequencing statistics

Additional file 1: Table S2. Total number of processed reads from High Iron (HI) and Low Iron (LI) experimental condition

Additional file 1: Table S3. Mapping results

Additional file 1: Table S4. Raw gene counts

Additional file 1: Figure S1. Sample to sample principal component analysis (PCA)

Additional file 2: Figure S2. *C. pseudotuberculosis* growth rate

Additional file 2: Figure S3. Effects of ethanol on bacterial growth

Additional file 3: Table S5. Differentially expressed genes identified in the T1 strain in response to iron limitation

Additional file 3: Table S6. Differentially expressed genes identified in the Cp13 mutant in response to iron limitation

Additional file 4: Table S7. Table of gene ontology terms assigned to DEGs identified in the T1 strain

Additional file 4: Table S8. Table of gene ontology terms assigned to DEGs identified in the Cp13 mutant

Additional file 5: Table S9. Predicted protein-protein interactions of iron regulated genes identified in the Cp13 mutant

Additional file 5: Table S10. Predicted protein-protein interactions of iron regulated genes identified in the Cp13 mutant

Additional file 5: Figure S4. Protein enrichment analysis

Additional file 6: Figure S5. Alignment of the amino acid sequence of the HtaA domain from protein products of the upregulated genes *htaC*, *htaA*, *htaF*, *htaG* and *Cp_3070* genes of *C. pseudotuberculosis*

Additional file 6: Figure S6. Structural protein domain characteristics of the HtaA, HtaC, HtaF, HtaG and Cp_3070 proteins

Additional file 6: Figure S7. Alignment of the amino acid sequences of the HmuU protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)

Additional file 6: Figure S8. Alignment of the amino acid sequences of the HmuT protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)

Additional file 6: Figure S9. Alignment of the amino acid sequences of the DNA-binding regulator *hrrA* of the two-component regulatory system *hrrSA* between *C. pseudotuberculosis* (Cp), *C. glutamicum* (Cg) and *C. diphtheriae* (Cd)

Additional file 7: Table S11. Genomic Island predictions

5.1.7. Abbreviations

ATP – Adenosine Triphosphate; BHI – Brain heart infusion medium; CFU – colony-forming unit; CLA - Caseous lymphadenitis; DEGs- differentially expressed genes; DIP – 2,2'-dipyridyl; EF – elongation factor; FDR – False discovery rate ; GO – Gene ontology; HI – High iron experimental condition; HMM – Hidden Markov Model; KEEG – Kyoto Encyclopedia of Genes and Genomes; LI – Low iron experimental condition; OD – Optical density; PCA - principal component analysis; PFAM – Protein Family database; PPI – protein-protein interaction; rlog – Regularize-logarithm transformed function; TCA – Tricarboxylic acid cycle or Krebs cycle; TF – Transcription factor; TRN – transcriptional regulatory network.

5.1.8. Acknowledgement

This research was financially supported by the National Council for Scientific and Technological Development (CNPq). We thank our collaborators from the Federal University of Para and the Federal University of Parana who provided important insight and expertise for the development of this project. We also would like to thank Dr. Emanuel Souza and Dra. Michelle Zibetti Tadra Sfeir from the biochemistry department of the University of Parana for performing the ion proton sequencing and for the excellent assistance in conducting the preparation of our RNA samples. We also like to acknowledge the sincere efforts from the members of the Laboratory of Cellular and Molecular Genetics (LGCM) in assisting the completion of this project.

5.1.9. Authors' contributions

Experimental design was previously determined by TC and AG. Growth and viability assays, transcriptomic sample preparation and RNA extraction were carried out by IC. Library preparation and sequencing were conducted by MZ under the supervision of ES. IC performed the bioinformatics analysis on the RNA-seq data. C. pseudotuberculosis regulatory network was done by MP and DP. RR assisted with the bioinformatics analysis. DI, PP, ARW and AG-N contributed substantially to the development and critically reviewed the manuscript. Data interpretation and revisions were conducted by IC and supervised by AG and VA.

5.1.10. Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationship that could construed as a potential conflict of interest.

5.1.11. Consent for publication

Not applicable.

5.1.12. Ethics approval and consent to participate

No ethics approval was required for any aspect of this study.

5.1.13. References

1. Baird GJ, Fontaine MC. Corynebacterium pseudotuberculosis and its Role in Ovine Caseous Lymphadenitis. J Comp Pathol. 2007;137(4):179–210.
2. Dorella FA, Carvalho Pacheco L, Oliveira SC, Miyoshi A, Azevedo V. Corynebacterium pseudotuberculosis : microbiology, biochemical properties, pathogenesis and molecular studies of virulence. Vet Res [Internet]. 2006 Mar;37(2):201–18. Available from: http://www.edpsciences.org/10.1051/vetres:2005056
3. Pépin M, Pardon P, Lantier F, Marly J, Levieux D, Lamand M. Experimental Corynebacterium pseudotuberculosis infection in lambs: kinetics of bacterial dissemination and inflammation. Vet Microbiol [Internet]. 1991 Feb;26(4):381–92. Available from: http://linkinghub.elsevier.com/retrieve/pii/037811359190031A
4. Cassat JE, Skaar EP. Iron in Infection and Immunity. Cell Host Microbe [Internet]. 2013 May;13(5):509–19. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1931312813001522
5. Schaible UE, Kaufmann SHE. Iron and microbial infection. Nat Rev Microbiol [Internet]. 2004 Dec 1;2(12):946–53. Available from: http://www.nature.com/articles/nrmicro1046
6. Schultz IJ, Chen C, Paw BH, Hamza I. Iron and Porphyrin Trafficking in Heme Biogenesis. J Biol Chem [Internet]. 2010 Aug 27;285(35):26753–9. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.R110.119503
7. Kehl-Fie TE, Skaar EP. Nutritional immunity beyond iron: a role for manganese and zinc. Curr Opin Chem Biol [Internet]. 2010 Apr;14(2):218–24. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1367593109001847
8. Weinberg ED. Microbial pathogens with impaired ability to acquire host iron. BioMetals. 2000;13(1):85–9.
9. Runyen-Janecky LJ. Role and regulation of heme iron acquisition in gram-negative pathogens. Front Cell Infect Microbiol [Internet]. 2013;3(October):1–11. Available from: http://journal.frontiersin.org/article/10.3389/fcimb.2013.00055/abstract
10. Sheldon JR, Laakso HA, Heinrichs DE. Iron acquisition strategies of bacterial pathogens. Virulence Mech Bact Pathog. 2016;(May 2016):In press.
11. Cornelis P, Wei Q, Andrews SC, Vinckx T. Iron homeostasis and management of oxidative stress response in bacteria. Met Integr Biometal Sci [Internet]. 2011;3(6):540–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21566833
12. Dorella FA, Estevam EM, Pacheco LGC, Guimaraes CT, Lana UGP, Gomes EA, et al. In Vivo Insertional Mutagenesis in Corynebacterium pseudotuberculosis: an Efficient Means To Identify DNA Sequences Encoding Exported Proteins. Appl Environ Microbiol [Internet]. 2006 Nov 1;72(11):7368–72. Available from: http://aem.asm.org/cgi/doi/10.1128/AEM.00294-06
13. Gibson CM, Caparon MG. Alkaline Phosphatase Reporter Transposon for Identification of Genes Encoding Secreted Proteins in Gram-Positive Microorganisms. Appl Environ Microbiol [Internet]. 2002 Feb 1;68(2):928–32. Available from: http://aem.asm.org/cgi/doi/10.1128/AEM.68.02.928-932.2002
14. Ling J, Pan H, Gao Q, Xiong L, Zhou Y, Zhang D, et al. Aerobactin Synthesis Genes iucA and iucC Contribute to the Pathogenicity of Avian Pathogenic

Escherichia coli O2 Strain E058. Pavelka M, editor. PLoS One [Internet]. 2013 Feb 27;8(2):e57794. Available from: http://dx.plos.org/10.1371/journal.pone.0057794
15. Kunkle CA, Schmitt MP. Analysis of a DtxR-Regulated Iron Transport and Siderophore Biosynthesis Gene Cluster in Corynebacterium diphtheriae. J Bacteriol [Internet]. 2005 Jan 15;187(2):422–33. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.187.2.422-433.2005
16. Billington SJ, Esmay PA, Songer J, Jost B, Hoppert M, Liebl W. Identification and role in virulence of putative iron acquisition genes from Corynebacterium pseudotuberculosis. FEMS Microbiol Lett. 2002;208(1):41–5.
17. Ribeiro D, Rocha F, Leite K, Soares S, Silva A, Portela R, et al. An iron-acquisition-deficient mutant of Corynebacterium pseudotuberculosis efficiently protects mice against challenge. Vet Res [Internet]. 2014;45(1):28. Available from: http://www.veterinaryresearch.org/content/45/1/28
18. 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 [Internet]. 2000;36(1):68–84. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10760164
19. Braun V. Iron uptake mechanisms and their regulation in pathogenic bacteria. Int J Med Microbiol [Internet]. 2001;291(2):67–79. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11437341&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1078/1438-4221-00103
20. Contreras H, Chim N, Credali A, Goulding CW. Heme uptake in bacterial pathogens. Curr Opin Chem Biol [Internet]. 2014;19(1):34–41. Available from: http://dx.doi.org/10.1016/j.cbpa.2013.12.014
21. Otto BR, Verweij-van Vught AMJJ, Maclaren DM. Transferrins and Heme-Compounds as Iron Sources for Pathogenic Bacteria. Crit Rev Microbiol [Internet]. 1992 Jan 25;18(3):217–33. Available from: http://www.tandfonline.com/doi/full/10.3109/10408419209114559
22. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial iron homeostasis. FEMS Microbiol Rev [Internet]. 2003;27(2–3):215–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12829269
23. Silva WM, Folador EL, Soares SC, Souza GHMF, Santos A V., Sousa CS, et al. Label-free quantitative proteomics of Corynebacterium pseudotuberculosis isolates reveals differences between Biovars ovis and equi strains. BMC Genomics. 2017;18(1):1–14.
24. Trost E, Götker S, Schneider J, Schneiker-Bekel S, Szczepanowski R, Tilker A, et al. Complete genome sequence and lifestyle of black-pigmented Corynebacterium aurimucosum ATCC 700975 (formerly C. nigricans CN-1) isolated from a vaginal swab of a woman with spontaneous abortion. BMC Genomics [Internet]. 2010;11:91. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20137072
25. Brune I, Werner H, Hüser AT, Kalinowski J, Pühler A, Tauch A. The DtxR protein acting as dual transcriptional regulator directs a global regulatory network involved in iron metabolism of Corynebacterium glutamicum. BMC Genomics. 2006;7:1–19.
26. Toyoda K, Teramoto H, Inui M, Yukawa H. Genome-Wide Identification of In Vivo Binding Sites of GlxR, a Cyclic AMP Receptor Protein-Type Regulator in Corynebacterium glutamicum. J Bacteriol [Internet]. 2011 Aug 15;193(16):4123–

33. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.00384-11
27. Brinkrolf K, Schröder J, Pühler A, Tauch A. The transcriptional regulatory repertoire of Corynebacterium glutamicum: Reconstruction of the network controlling pathways involved in lysine and glutamate production. J Biotechnol [Internet]. 2010 Sep 1;149(3):173–82. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0168165609005549
28. Skaar EP. Iron-Source Preference of Staphylococcus aureus Infections. Science (80-) [Internet]. 2004 Sep 10;305(5690):1626–8. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1099930
29. Allen CE, Schmitt MP. Novel Hemin Binding Domains in the Corynebacterium diphtheriae HtaA Protein Interact with Hemoglobin and Are Critical for Heme Iron Utilization by HtaA. J Bacteriol [Internet]. 2011 Oct 1;193(19):5374–85. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.05508-11
30. Allen CE, Burgos JM, Schmitt MP. Analysis of novel iron-regulated, surface-anchored hemin-binding proteins in Corynebacterium diphtheriae. J Bacteriol [Internet]. 2013;195(12):2852–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23585541
31. Popovic T, Kombarova SY, Reeves MW, Nakao H, Mazurova IK, Wharton M, et al. Molecular epidemiology of diphtheria in Russia, 1985-1994. J Infect Dis [Internet]. 1996 Nov;174(5):1064–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8896510
32. Allen CE, Schmitt MP. HtaA is an iron-regulated hemin binding protein involved in the utilization of heme iron in Corynebacterium diphtheriae. J Bacteriol [Internet]. 2009;191(8):2638–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19201805
33. Kumar R, Lovell S, Matsumura H, Battaile KP, Moënne-Loccoz P, Rivera M. The Hemophore HasA from Yersinia pestis (HasA yp) Coordinates Hemin with a Single Residue, Tyr75, and with Minimal Conformational Change. Biochemistry [Internet]. 2013 Apr 23;52(16):2705–7. Available from: http://pubs.acs.org/doi/10.1021/bi400280z
34. Honsa ES, Owens CP, Goulding CW, Maresso AW. The Near-iron Transporter (NEAT) Domains of the Anthrax Hemophore IsdX2 Require a Critical Glutamine to Extract Heme from Methemoglobin. J Biol Chem [Internet]. 2013 Mar 22;288(12):8479–90. Available from: http://www.jbc.org/lookup/doi/10.1074/jbc.M112.430009
35. Ekworomadu MT, Poor CB, Owens CP, Balderas MA, Fabian M, Olson JS, et al. Differential Function of Lip Residues in the Mechanism and Biology of an Anthrax Hemophore. Isberg RR, editor. PLoS Pathog [Internet]. 2012 Mar 8;8(3):e1002559. Available from: http://dx.plos.org/10.1371/journal.ppat.1002559
36. Mokry DZ, Nadia-Albete A, Johnson MK, Lukat-Rodgers GS, Rodgers KR, Lanzilotta WN. Spectroscopic evidence for a 5-coordinate oxygenic ligated high spin ferric heme moiety in the Neisseria meningitidis hemoglobin binding receptor. Biochim Biophys Acta - Gen Subj [Internet]. 2014 Oct;1840(10):3058–66. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0304416514002293
37. Bibb LA, Kunkle CA, Schmitt MP. The ChrA-ChrS and HrrA-HrrS Signal Transduction Systems Are Required for Activation of the hmuO Promoter and Repression of the hemA Promoter in Corynebacterium diphtheriae. Infect Immun [Internet]. 2007 May 1;75(5):2421–31. Available from:

http://iai.asm.org/cgi/doi/10.1128/IAI.01821-06
38. Kunkle CA, Schmitt MP. Analysis of the Corynebacterium diphtheriae DtxR Regulon: Identification of a Putative Siderophore Synthesis and Transport System That Is Similar to the Yersinia High-Pathogenicity Island-Encoded Yersiniabactin Synthesis and Uptake System. J Bacteriol [Internet]. 2003 Dec 1;185(23):6826–40. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.185.23.6826-6840.2003
39. Silva WM, Carvalho RDDO, Dorella FA, Folador EL, Souza GHMF, Pimenta AMC, et al. Quantitative Proteomic Analysis Reveals Changes in the Benchmark Corynebacterium pseudotuberculosis Biovar Equi Exoproteome after Passage in a Murine Host. Front Cell Infect Microbiol [Internet]. 2017;7(July). Available from: http://journal.frontiersin.org/article/10.3389/fcimb.2017.00325/full
40. Kao WC, Kleinschroth T, Nitschke W, Baymann F, Neehaul Y, Hellwig P, et al. The obligate respiratory supercomplex from Actinobacteria. Biochim Biophys Acta - Bioenerg [Internet]. 2016;1857(10):1705–14. Available from: http://dx.doi.org/10.1016/j.bbabio.2016.07.009
41. Sone N, Nagata K, Kojima H, Tajima J, Kodera Y, Kanamaru T, et al. A novel hydrophobic diheme c-type cytochrome. Purification from Corynebacterium glutamicum and analysis of the QcrCBA operon encoding three subunit proteins of a putative cytochrome reductase complex. Biochim Biophys Acta [Internet]. 2001;1503(3):279–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11115640
42. Bott M, Niebisch A. The respiratory chain of Corynebacterium glutamicum. J Biotechnol. 2003;104(1–3):129–53.
43. Park SJ, Tseng CP, Gunsalus RP. Regulation of succinate dehydrogenase (sdhCDAB) operon expression in Escherichia coli in response to carbon supply and anaerobiosis: role of ArcA and Fnr. Mol Microbiol [Internet]. 1995 Feb;15(3):473–82. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7783618
44. Lancaster CR, Kröger A. Succinate: quinone oxidoreductases: new insights from X-ray crystal structures. Biochim Biophys Acta [Internet]. 2000 Aug 15;1459(2–3):422–31. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11004459
45. Oexle H, Gnaiger E, Weiss G. Iron-dependent changes in cellular energy metabolism: influence on citric acid cycle and oxidative phosphorylation. Biochim Biophys Acta [Internet]. 1999 Nov 10;1413(3):99–107. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10556622
46. Roset MS, Alefantis TG, DelVecchio VG, Briones G. Iron-dependent reconfiguration of the proteome underlies the intracellular lifestyle of Brucella abortus. Sci Rep [Internet]. 2017 Dec 6;7(1):10637. Available from: http://www.nature.com/articles/s41598-017-11283-0
47. Iqbal I, Bajeli S, Akela A, Kumar A. Bioenergetics of Mycobacterium: An Emerging Landscape for Drug Discovery. Pathogens [Internet]. 2018 Feb 23;7(1):24. Available from: http://www.mdpi.com/2076-0817/7/1/24
48. Poole RK, Cook GM. Redundancy of aerobic respiratory chains in bacteria? Routes, reasons and regulation. 2000;165–224. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0065291100430055
49. Giuffrè A, Borisov VB, Arese M, Sarti P, Forte E. Cytochrome bd oxidase and bacterial tolerance to oxidative and nitrosative stress. Biochim Biophys Acta [Internet]. 2014 Jul;1837(7):1178–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24486503

50. Charbon G, Campion C, Chan SHJ, Bjørn L, Weimann A, da Silva LCN, et al. Re-wiring of energy metabolism promotes viability during hyperreplication stress in E. coli. Walker GC, editor. PLOS Genet [Internet]. 2017 Jan 27;13(1):e1006590. Available from: http://dx.plos.org/10.1371/journal.pgen.1006590
51. Pontes MH, Sevostyanova A, Groisman EA. When Too Much ATP Is Bad for Protein Synthesis. J Mol Biol [Internet]. 2015 Aug;427(16):2586–94. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022283615003617
52. Skerker JM, Prasol MS, Perchuk BS, Biondi EG, Laub MT. Two-Component Signal Transduction Pathways Regulating Growth and Cell Cycle Progression in a Bacterium: A System-Level Analysis. Arkin A, editor. PLoS Biol [Internet]. 2005 Sep 27;3(10):e334. Available from: http://dx.plos.org/10.1371/journal.pbio.0030334
53. Frunzke J, Gätgens C, Brocker M, Bott M. Control of heme homeostasis in Corynebacterium glutamicum by the two-component system HrrSA. J Bacteriol. 2011;193(5):1212–21.
54. Wennerhold J, Bott M. The DtxR Regulon of Corynebacterium glutamicum. J Bacteriol [Internet]. 2006;188(8):2907–18. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1446976/
55. Wennerhold J, Krug A, Bott M. The AraC-type Regulator RipA Represses Aconitase and Other Iron Proteins from Corynebacterium under Iron Limitation and Is Itself Repressed by DtxR. J Biol Chem [Internet]. 2005;280(49):40500–8. Available from: http://www.jbc.org/content/280/49/40500
56. Novichkov PS, Kazakov AE, Ravcheev DA, Leyn SA, Kovaleva GY, Sutormin RA, et al. RegPrecise 3.0 – A resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genomics [Internet]. 2013;14(1):745. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-745
57. Rickman L, Scott C, Hunt DM, Hutchinson T, Menéndez MC, Whalan R, et al. A member of the cAMP receptor protein family of transcription regulators in Mycobacterium tuberculosis is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. Mol Microbiol [Internet]. 2005 Mar 30;56(5):1274–86. Available from: http://doi.wiley.com/10.1111/j.1365-2958.2005.04609.x
58. Choi W-W, Park S-D, Lee S-M, Kim H-B, Kim Y, Lee H-S. The whcA gene plays a negative role in oxidative stress response of Corynebacterium glutamicum. FEMS Microbiol Lett [Internet]. 2009 Jan;290(1):32–8. Available from: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2008.01398.x
59. Alam MS, Garg SK, Agrawal P. Studies on structural and functional divergence among seven WhiB proteins of Mycobacterium tuberculosis H37Rv. FEBS J [Internet]. 2009 Jan;276(1):76–93. Available from: http://doi.wiley.com/10.1111/j.1742-4658.2008.06755.x
60. Wu J, Ru H, Xiang Z, Jiang J, Wang Y, Zhang L, et al. WhiB4 Regulates the PE/PPE Gene Family and is Essential for Virulence of Mycobacterium marinum. Sci Rep [Internet]. 2017 Dec 7;7(1):3007. Available from: http://www.nature.com/articles/s41598-017-03020-4
61. Kaprelyants AS, Mukamolova G V, Ruggiero A, Makarov VA, Demina GR, Shleeva MO, et al. Resuscitation-promoting Factors (Rpf): In Search of Inhibitors [Internet]. Protein & Peptide Letters. 2012. Available from:

http://www.eurekaselect.com/102449/article
62. Telkov M V, Demina GR, Voloshin SA, Salina EG, Dudik T V, Stekhanova TN, et al. Proteins of the Rpf (resuscitation promoting factor) family are peptidoglycan hydrolases. Biochemistry (Mosc) [Internet]. 2006 Apr;71(4):414–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16615861
63. Jungwirth B, Emer D, Brune I, Hansmeier N, PÃ¼hler A, Eikmanns BJ, et al. Triple transcriptional control of the resuscitation promoting factor 2 (rpf2) gene of Corynebacterium glutamicum by the regulators of acetate metabolism RamA and RamB and the cAMP-dependent regulator GlxR. FEMS Microbiol Lett [Internet]. 2008 Apr;281(2):190–7. Available from: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2008.01098.x
64. Kana BD, Mizrahi V. Resuscitation-promoting factors as lytic enzymes for bacterial growth and signaling. FEMS Immunol Med Microbiol. 2010;58(1):39–50.
65. Sheldon JR, Heinrichs DE. Recent developments in understanding the iron acquisition strategies of gram positive pathogens. FEMS Microbiol Rev. 2015;39(4):592–630.
66. Braun V, Herrmann C. Docking of the Periplasmic FecB Binding Protein to the FecCD Transmembrane Proteins in the Ferric Citrate Transport System of Escherichia coli. J Bacteriol [Internet]. 2007 Oct 1;189(19):6913–8. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.00884-07
67. Bibb LA, King ND, Kunkle CA, Schmitt MP. Analysis of a Heme-Dependent Signal Transduction System in Corynebacterium diphtheriae: Deletion of the chrAS Genes Results in Heme Sensitivity and Diminished Heme-Dependent Activation of the hmuO Promoter. Infect Immun [Internet]. 2005 Nov 1;73(11):7406–12. Available from: http://iai.asm.org/cgi/doi/10.1128/IAI.73.11.7406-7412.2005
68. Kunkle CA, Schmitt MP. Comparative Analysis of hmuO Function and Expression in Corynebacterium Species. J Bacteriol [Internet]. 2007 May 1;189(9):3650–4. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.00056-07
69. Baumbach J, Rahmann S, Tauch A. Reliable transfer of transcriptional gene regulatory networks between taxonomically related organisms. BMC Syst Biol [Internet]. 2009;3(1):8. Available from: http://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-3-8
70. Qian Y, Lee JH, Holmes RK. Identification of a DtxR-Regulated Operon That Is Essential for Siderophore-Dependent Iron Uptake in Corynebacterium diphtheriae. J Bacteriol [Internet]. 2002 Sep 1;184(17):4846–56. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.184.17.4846-4856.2002
71. Brune I, Brinkrolf K, Kalinowski J, Pühler A, Tauch A. The individual and common repertoire of DNA-binding transcriptional regulators of Corynebacterium glutamicum, Corynebacterium efficiens, Corynebacterium diphtheriae and Corynebacterium jeikeium deduced from the complete genome sequences. BMC Genomics [Internet]. 2005 Jun 7;6:86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15938759
72. Merchant AT, Spatafora GA. A role for the DtxR family of metalloregulators in gram-positive pathogenesis. Mol Oral Microbiol [Internet]. 2014 Feb;29(1):1–10. Available from: http://doi.wiley.com/10.1111/omi.12039
73. Muhl D, Jessberger N, Hasselt K, Jardin C, Sticht H, Burkovski A. DNA binding by Corynebacterium glutamicum TetR-type transcription regulator AmtR. BMC

Mol Biol [Internet]. 2009 Jul 23;10:73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19627583
74. Shea CM, McIntosh MA. Nucleotide sequence and genetic organization of the ferric enterobactin transport system: homology to other periplasmic binding protein-dependent systems in Escherichia coli. Mol Microbiol [Internet]. 1991 Jun;5(6):1415–28. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1838574
75. 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 [Internet]. 2009 May 1;191(9):2964–72. Available from: http://jb.asm.org/cgi/doi/10.1128/JB.00112-09
76. Schröder J, Tauch A. Transcriptional regulation of gene expression in Corynebacterium glutamicum: The role of global, master and local regulators in the modular and hierarchical gene regulatory network. FEMS Microbiol Rev. 2010;34(5):685–737.
77. Andrews S. FastQC: a quality control tool for high throughput sequence data [Internet]. 2010 [cited 2018 Aug 20]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
78. Williams CR, Baccarella A, Parrish JZ, Kim CC. Trimming of sequence reads alters RNA-Seq gene expression estimates. BMC Bioinformatics [Internet]. 2016 Dec 25;17(1):103. Available from: http://www.biomedcentral.com/1471-2105/17/103
79. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics [Internet]. 2014 Aug 1;30(15):2114–20. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu170
80. Yuan Y, Xu H, Leung RK-K. An optimized protocol for generation and analysis of Ion Proton sequencing reads for RNA-Seq. BMC Genomics [Internet]. 2016 Dec 26;17(1):403. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2745-8
81. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics [Internet]. 2015 Jan 15;31(2):166–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu638
82. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.
83. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res [Internet]. 2017 Jan 4;45(D1):D362–8. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw937
84. Röttger R, Kalaghatgi P, Sun P, Soares S de C, Azevedo V, Wittkop T, et al. Density parameter estimation for finding clusters of homologous proteins—tracing actinobacterial pathogenicity lifestyles. Bioinformatics [Internet]. 2013 Jan 15;29(2):215–22. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bts653
85. Wheeler TJ, Eddy SR. nhmmer: DNA homology search with profile HMMs. Bioinformatics [Internet]. 2013 Oct 1;29(19):2487–9. Available from:

https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btt403
86. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res [Internet]. 2011 Jul 1;39(suppl):W29–37. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr367
87. Soares SC, Geyik H, Ramos RTJ, de Sá PHCG, Barbosa EGV, Baumbach J, et al. GIPSy: Genomic island prediction software. J Biotechnol [Internet]. 2016 Aug;232:2–11. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168165615301152
88. Society RS. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by : . 2016;57(1):289–300.

Table 1: Differentially expressed genes identified in the T1 wild-type and Cp13 mutant strain of *C. pseudotuberculosis.*

Gene/Locus ID	**Product**	**LI:HI Fold Change**		***p*-value***		**Gene reg.**	
		T1	**Cp13**	**T1**	**Cp13**	**T1**	**Cp13**
Cp1002B_95	Hypothetical protein	1.70	1.45	0.000814	0.019094	DOWN	UNC.
sdcS	Sodium-dependent dicarboxylate transporter	2.35	1.47	6.21E-16	0.011373	DOWN	UNC.
Cp1002B_1235	Hypothetical protein	2.17	1.99	0.002796	0.000258	DOWN	DOWN
ccsA	Cytochrome c biogenesis protein ccsa	1.53	1.92	0.022829	0.002895	UP	UP
htaA	Cell-surface hemin receptor	1.95	2.70	2.66E-12	0.000128	UP	UP
hmuT	ABC transporter substrate-binding protein	1.86	3.04	1.19E-11	5.47E-05	UP	UP
hmuU	Hemin import atp-binding protein	1.50	2.09	3.31E-05	0.005646	UP	UP
htaC	Hypothetical protein	6.93	7.41	9.8E-66	3.33E-11	UP	UP
Cp1002B_2935	Transglycosylase associated protein	2.47	1.75	6.46E-15	3.39E-05	DOWN	DOWN
hrrA	DNA-binding response regulator	1.68	1.99	3.02E-06	4.02E-05	UP	UP
Cp1002B_3070	Hypothetical protein	8.49	22.90	1.74E-34	3.62E-26	UP	UP
Cp1002B_3075	Hypothetical protein	10.19	28.20	1.71E-47	2.43E-32	UP	UP
Cp1002B_3725	FUSC family protein	1.34	1.51	0.004667	2.37E-05	UNC.	UP
Cp1002B_4550	Antimicrobial peptide ABC transporter	1.55	1.40	0.003592	0.003453	DOWN	UNC.
htaF	Cell-surface hemin receptor	3.38	7.29	4.16E-27	2.69E-14	UP	UP
htaG	Uncharacterized protein htac	2.78	5.86	3.62E-24	0.002598	UP	UP
ctaC	Cytochrome c oxidase subunit II	2.48	2.10	5.17E-13	1.06E-06	DOWN	DOWN
ctaF	Cytochrome c oxidase polypeptide 4	2.23	2.13	7.95E-11	1.08E-07	DOWN	DOWN
ctaE	Cytochrome c oxidase subunit III	2.31	2.61	9.09E-10	7.24E-11	DOWN	DOWN
qcrC	Ubiquinol-cytochrome C reductase cytochrome C	2.84	2.57	1.19E-14	1.02E-09	DOWN	DOWN
qcrA	Ubiquinol-cytochrome c reductase iron-sulfur	2.35	2.18	7.6E-05	0.000568	DOWN	DOWN
qcrB	Ubiquinol-cytochrome C reductase cytochrome B	2.42	2.44	4.7E-13	6.48E-07	DOWN	DOWN
Cp1002B_5915	Hypothetical protein	1.43	2.10	0.006998	1.59E-11	UNC.	UP

rbfA	Ribosome-binding factor A	1.62	1.54	0.00192	0.002598	DOWN	DOWN
Cp1002B_8515	Hypothetical protein	1.33	1.55	0.039677	0.01409	UNC.	UP
ywlC	Trna threonylcarbamoyladenosine biosynthesis protein ywlc	1.60	1.52	0.016627	3.39E-05	UP	UP
ripA	HTH-type transcriptional repressor of iron protein A	2.13	2.44	1.193E-12	6.07E-13	UP	UP
rpfB	Resuscitation-promoting factor rpfb	1.32	1.70	0.039677	1.58E-07	UNC.	UP
fkbP	Peptidyl-prolyl cis-trans isomerase, FKBP-type	1.97	1.61	0.000242	0.013233	DOWN	DOWN
Cp1002B_10210	Substrate-binding protein	5.43	2.44	2.89E-17	1.28E-07	DOWN	DOWN
Cp1002B_10495	Hypothetical protein	4.97	6.30	3.14E-62	2.66E-14	UP	UP
Cp1002B_10785	Hypothetical protein	3.59	5.93	2.63E-17	3.23E-32	UP	UP

Fold change values correspond to the average normalized count ratio between LI (low iron) media and HI (high iron) media. **p*-values were adjusted using the Benjamin and Hochberg false discovery rate (FDR) approach. Genes with *p*-values < 0.05 and an absolute FC >1.5 were considered differentially expressed. Gene regulation is based on the LI:HI ratio expression value: UP, up-regulated; UNC. unchanged (FC<1.5); DOWN, down-regulated.

5.1.14. Figure List:

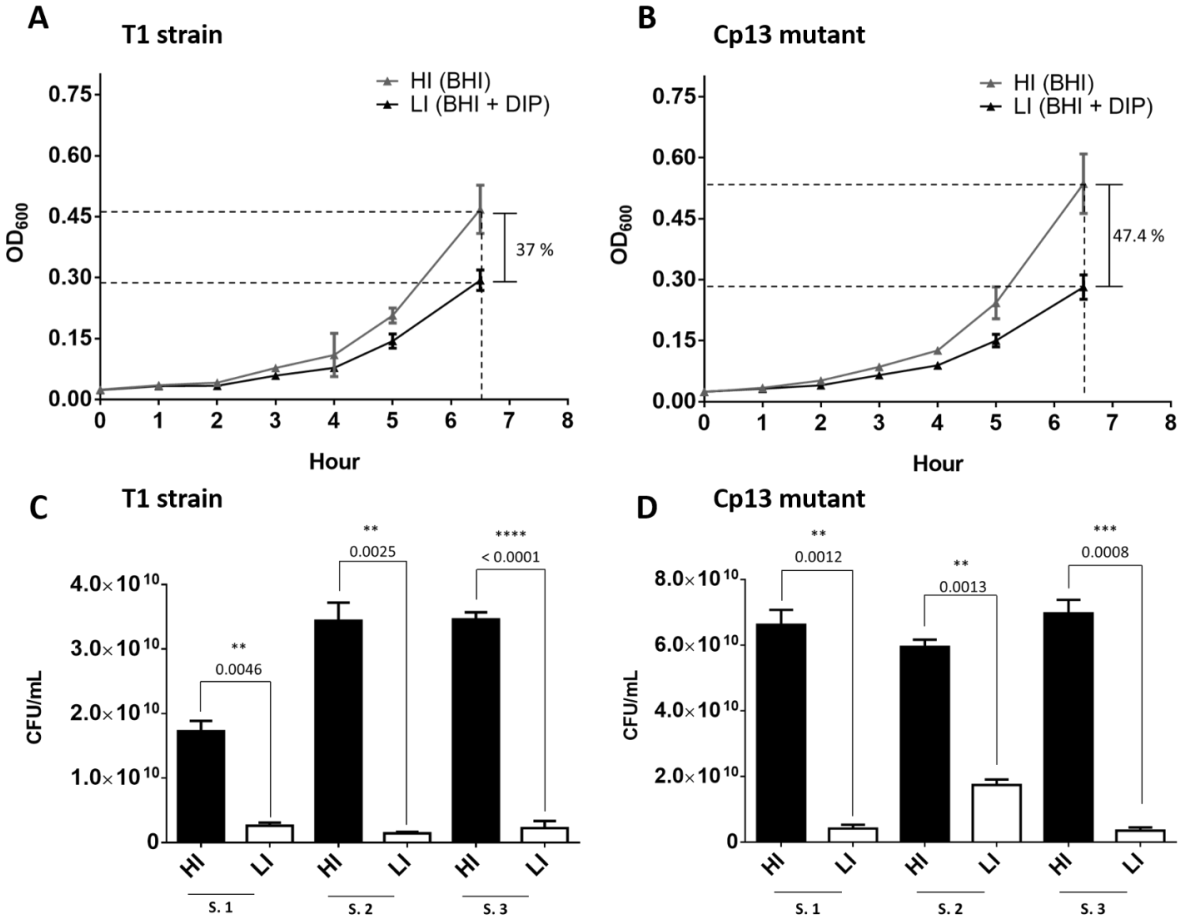


Figure 3. *C. pseudotuberculosis* growth assays.

Growth proliferation was determined by measuring the OD of low iron (LI) and high iron (HI) cultures at 600nm of the parental T1 strain and the Cp13 mutant strain, reflecting the growth dynamics of the bacterial strains in the culture media over the different time points. BHI broth was supplemented with 250 μM of the iron chelator 2,2'-dipyridyl (DIP) to iron stress the bacteria. The stress was confirmed at 6h30min with a significant reduction in the final culture density of both strains. Symbols indicate means ± standard deviation from triplicate determinants. At the 6h30min, both strains presented a final LI culture density of 63% and 53% of the bacteria grown in the high iron BHI broth (**A**) parental T1 strain and (**B**) Cp13 mutant. *C. pseudotuberculosis* viability was measured by determining the number of CFU/mL after 6h30min of incubation in a low iron (LI) dipyridyl-chelated and high iron (HI) non-chelated media. CFU counts are expressed by means ± s. d. (n= 3 for each of the 3 samples) and a paired t-test was used for two-group comparison (LI vs HI). A p value < 0.05 was considered significant. (**C**) CFU in the wild-type T1 strain (**D**) and the Cp13 mutant

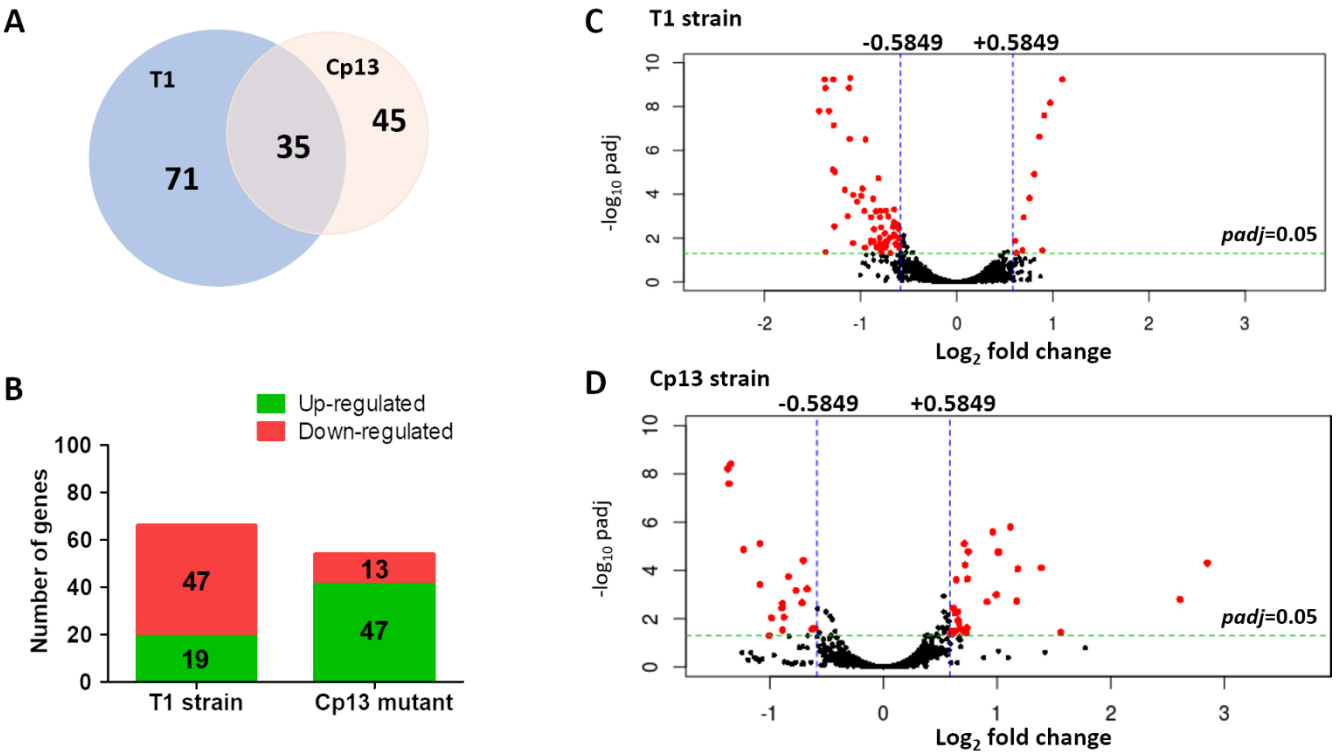


Figure 4. Differential gene expression of *C. pseudotuberculosis* Cp13 mutant and T1 wild-type under iron limitation (LI/HI).

All genes with an adjusted *p* value < 0.05 (adjusted for multiple comparisons by Benjamin-Hochberg procedure [83]) were considered differentially expressed. (**A**) Scaled Venn representing the expression of 106 DE genes identified in the T1 strain and 80 genes identified in the Cp13 mutant under iron limitation. 35 DEGs were commonly expressed in the T1 strain and the Cp13 mutant. A stacked-bar graph representing the total number of the up and down-regulated genes identified in the T1 and Cp13 strains, considering a greater than 1.5-fold change (log2fold change 0.5849 and -0.5849). (**B**) 19 genes up-regulated and 47 down-regulated in T1 strain. (**C**) 41 genes were up-regulated and 13 were down-regulated in the Cp13 mutant. Volcano plot of the log2-fold change of each detected gene in relation to their -log10 of adjusted *p* values. Threshold of log2fold change and *p* adjusted value are represented by blue and green lines, respectively. Volcano plot representation of DE genes identified in the: (**D**) T1 parental strain and (**E**) Cp13 mutant strain.

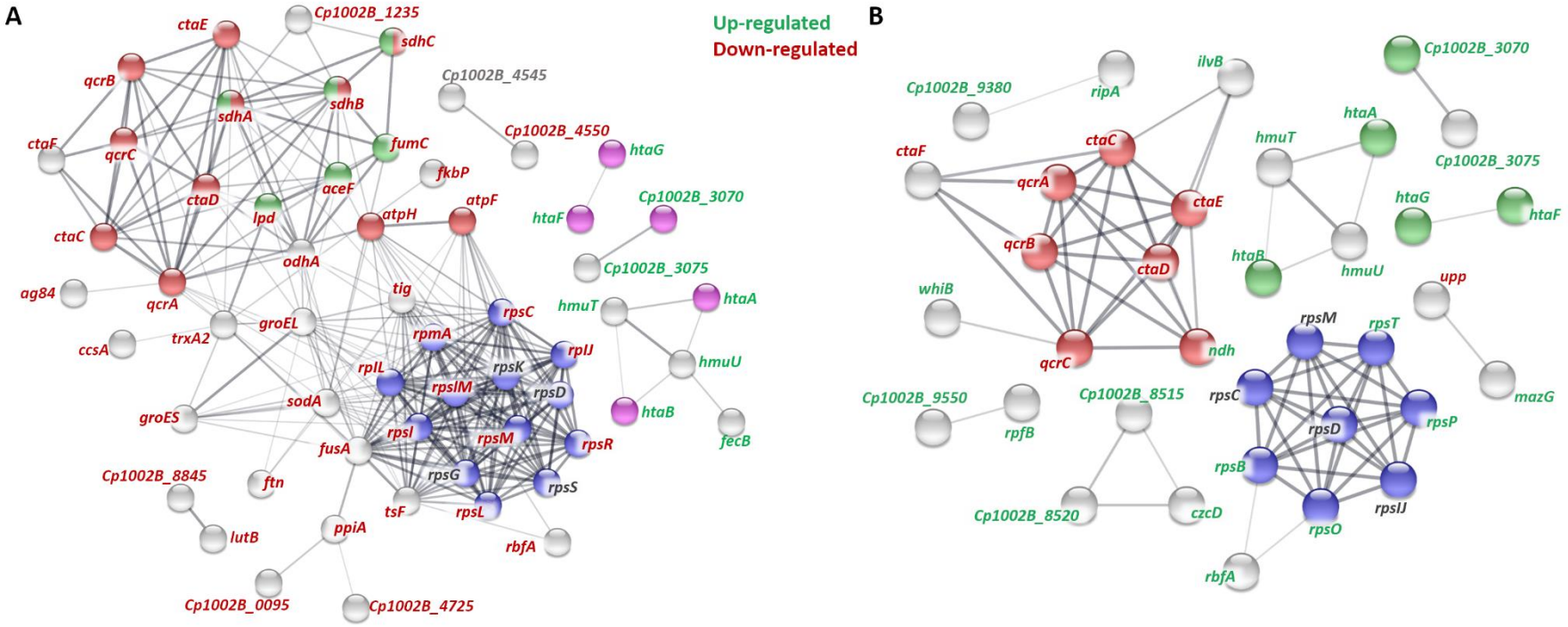


Figure 5. Protein interaction network analyses of up and down-regulated genes in *C. pseudotuberculosis* strains T1 and Cp13.

PPI analyses were carried out using the STRING database analysis tool and line thickness indicates strength of data support for each interaction. Connected nodes and interaction with a medium (>0.4), high (>0.7) and highest confidence (>0.9) are visualized in the network. Node colors represent enriched pathway and domains identified in both strains and gene identification color represents up-regulation (red), down-regulation (blue) and unchanged expression (gray). *p*-value of PPI interactions indicates significance of protein association. The line thickness represented in the networks indicates the strength of data support for each interaction. (**A**) The STRING PPI network of the T1 strain contained 74 nodes and 286 edges with PPI interaction enrichment *p*-value of <1.0e-16. 4 enriched pathways are shown: TCA cycle (green), oxidative phosphorylation (red), ribosome (blue) and HtaA domain (pink). (**B**) Cp13 PPI interactions contained 57 nodes and 73 edges with a PPI interaction enrichment *p*-value of 1.27e-11. 3 enriched pathways are shown: oxidative phosphorylation (red), ribosome (blue) and HtaA domain (green). Pathway enrichment based on p *adjusted* values after Benjamini-Hochberg discovery rate correction (Benjamini-Hochberg, 1995). Complete PPI list for each of the strains are provided in Additional file 5.

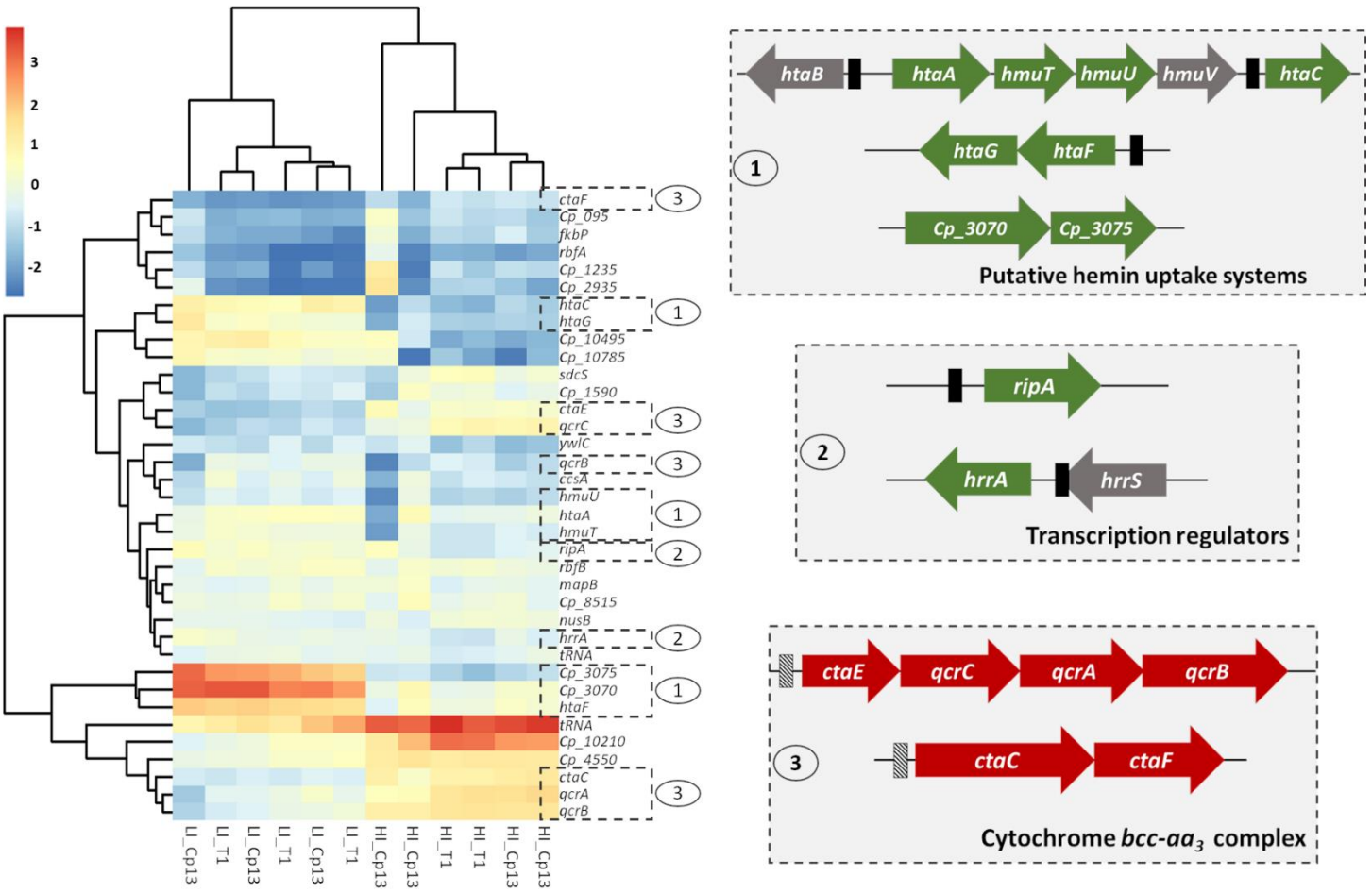


Figure 6. Expression pattern of commonly expressed genes identified in the T1 strain and the Cp13 mutant.

A hierarchical clustering of the 35 DEGs commonly expressed between the strains was used to identify the pattern of expression of these genes. Clustering of normalized DESEq2 count data show per gene log2 deviation from gene's average across all samples using *rlog* transformed data and revels 25 DEGs consistently up- or down-regulated in the T1 strain and Cp13 mutant. Columns represent sample identification in relation to experimental condition and strains. Representative gene clusters identified between both strains are indicated highlighting the clusters involving (1) high-affinity hemin-binding acquisition systems, (2) transcription factors and genes encoding the (3) cytochrome *bcc-aa3* super complex of the respiratory chain. Up-regulated genes are shown as green arrows; down-regulated genes are shown as red arrows and genes with unchanged expression are shown as gray. Black solid bars represent DtxR predicted regulatory binding sites and striped-bars represent GlxR regulatory binding sites.

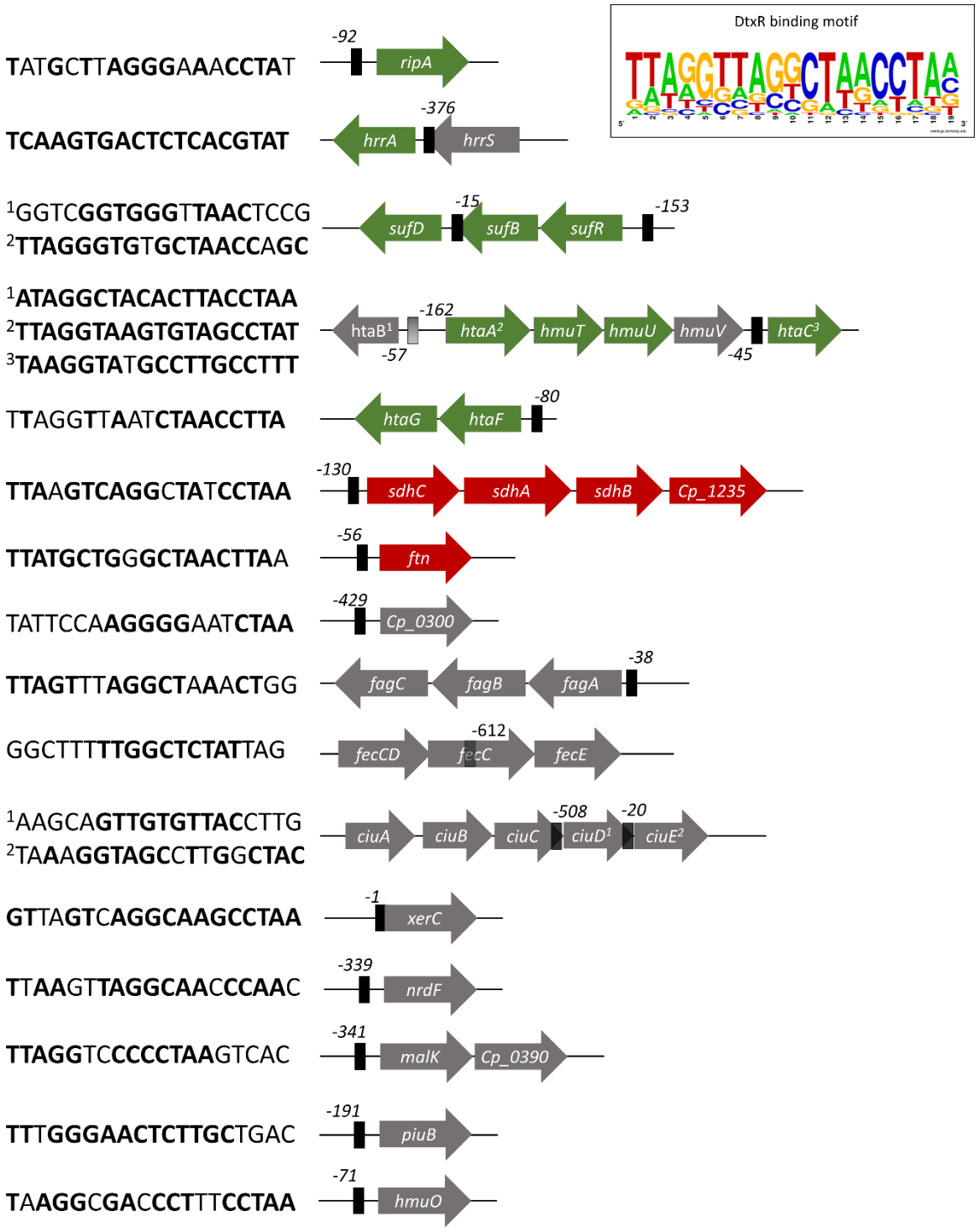


Figure 7. Genetic map of the target genes identified as having a putative iron-DtxR regulated binding site.

Conserved residues in relation to the reference are indicate in bold. DtxR target site prediction was based on orthologous target and experimentally confirmed sites of DtxR orthologous identified in taxonomically closely related species: *C. glutamicum* (NC_003450), *C. diphtheriae* (*NZ_LN831026*) *and M. tuberculosis* (NC_000962). Upregulated genes are shown as green arrows; downregulated genes are shown as red arrows and genes with unchanged expression are shown as gray. Detected DtxR binding sites are indicated as black boxes upstream the genes coding regions. The 19-bp consensus sequence of the DtxR-binding site of *C. pseudotuberculosis* is shown as DNA sequence logo.

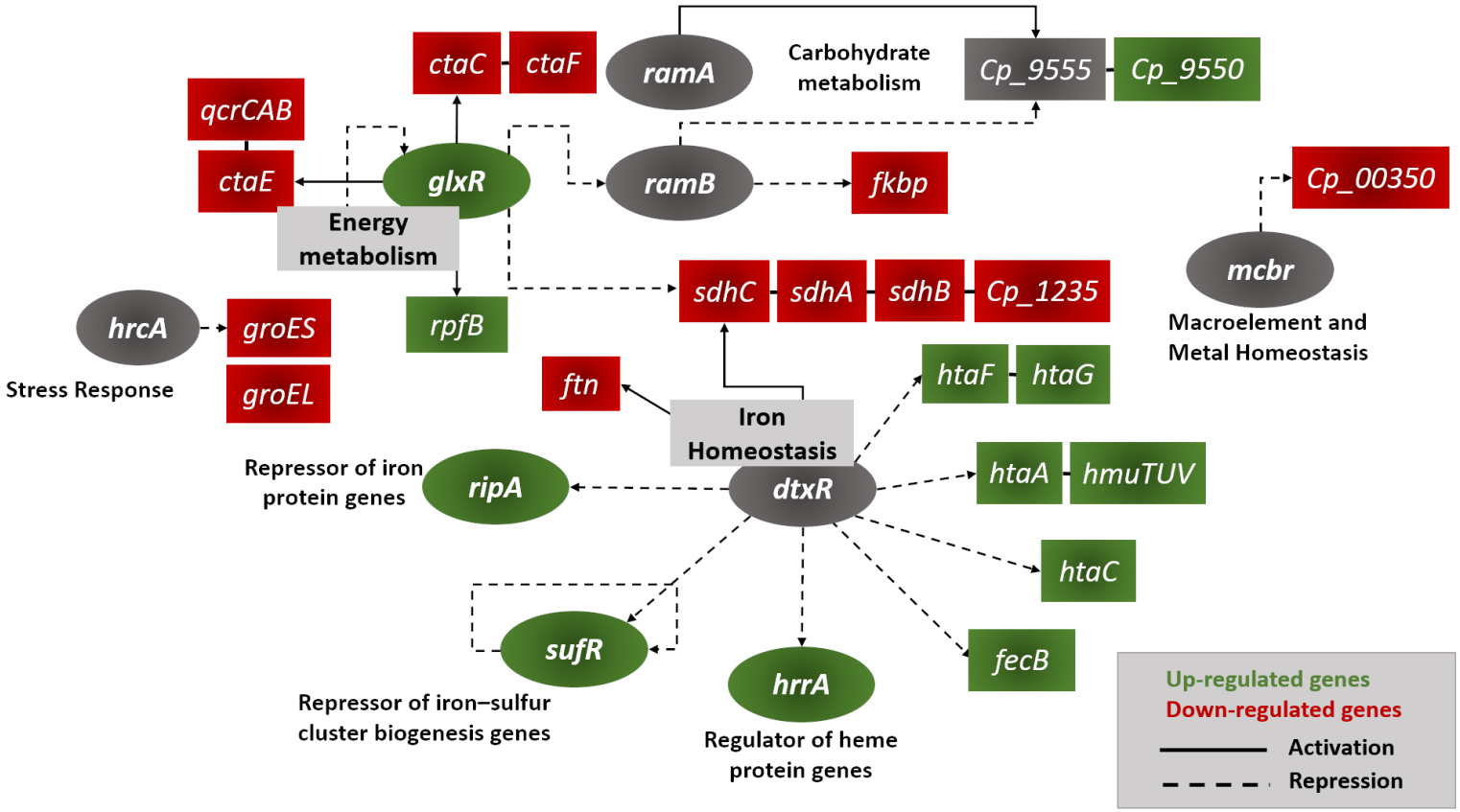


Figure 8. ***C. pseudotuberculosis*** **transcriptional regulatory iron network**

Networks shown involve DEGs identified in the transcriptional profile of both strains. TF genes are shown as circles. The colors represent the type of expression pattern and it is consistent with data in Table 1. Green indicates up-regulated genes, red indicates down-regulated genes and gray indicates genes with no difference in fold-change expression. Black solid lines represent transcriptional activation of the target gene and black dashed lines represent transcriptional repression of the target gene. Connected genes represent opero**NS.**

5.2. Resultados Adicionais

Nessa seção serão apresentados resultados e discussão que não foram relatados no manuscrito (subtópico 4.1 dessa seção) ou informações adicionais pertinentes a realização do projeto. Em sua maioria, esses resultados incluem as etapas de padronização, controle de qualidade das amostras extraídas e estatística dos dados de sequenciamento. Em adição, algumas tabelas apresentadas no material suplementar do manuscrito também foram incluídas, uma vez que vimos a necessidade de discutir uma pouco mais esses resultados. Esses resultados permitiram a viabilização do projeto, mas não são incluídos no manuscrito final devido a uma conjuntura de fatores, dentre eles a limitação de espaço destinados a publicação.

5.2.1. Quantificação da concentração de ferro no meio BHI e avaliação do crescimento das linhagens T1 parental e Cp13 mutante de C. pseudotuberculosis em meio BHI

O meio BHI é rotineiramente utilizado para o cultivo de *C. pseudotuberculosis*. Esse meio é extramente rico, contudo sua exata constituição não é conhecida. Como o objetivo geral deste trabalho envolve a resposta transcricional bacteriana em relação a restrição de ferro, os níveis de ferro presentes no meio BHI foram determinados utilizando a espectrometria de absorção atômica com chama. O meio BHI dosado apresentou uma concentração de 0,326 µg/mL (5,8411 µM) de ferro, concentração 3 vezes maior que a concentração necessária para o crescimento microbiano (HARTMANN; BRAUN, 1981; SHELDON; LAAKSO; HEINRICHS, 2016).

A fim de avaliar se a mutação do gene *ciuA* afetaria o crescimento da bactéria, curvas de crescimento foram realizadas com as linhagens tipo selvagem (T1) e mutante (Cp13) em meio BHI rico em ferro (Fig. 9). O perfil de crescimento das duas linhagens foi muito similar no meio HI, sendo que ambas as linhagens iniciaram a fase exponencial ao mesmo tempo e apresentaram taxa de crescimento similar ao longo da incubação. A viabilidade dessas duas linhagens também foi avaliada, e novamente, não foram observadas diferenças significativas entre o número de UFC entre as duas linhagens ($p>0.5$).

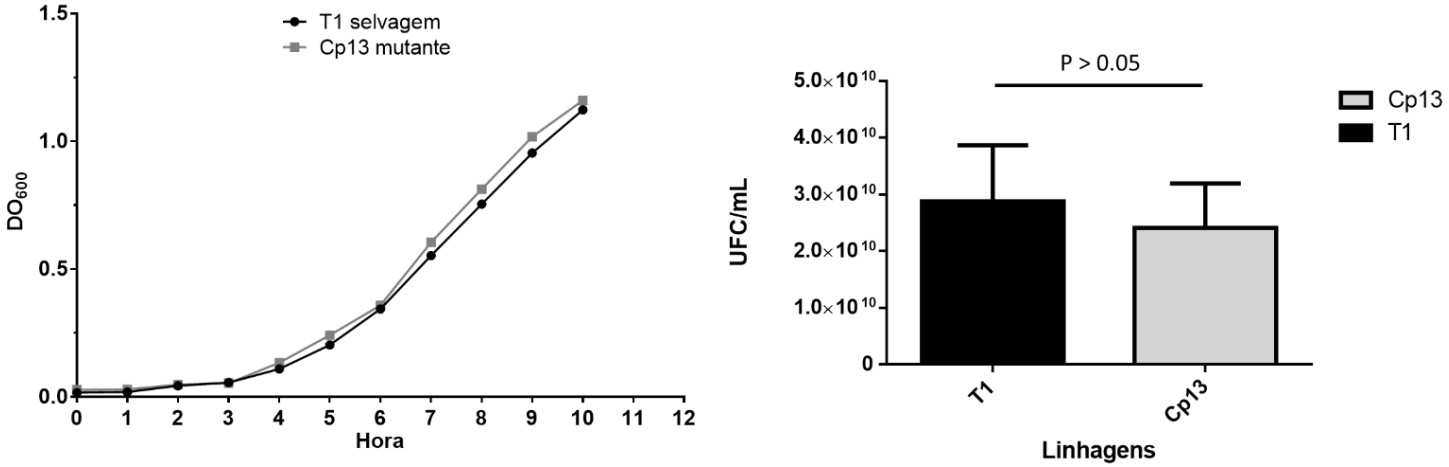


Figure 9. Curvas de crescimento e viabilidade bacteriana das linhagens selvagem (T1) e mutante (Cp13) em meio BHI.

(A) Culturas de *C. pseudotuberculosis* crescidas *overnight* em meio líquido foram diluídas a uma densidade ótica inicial de 0.02 em meio BHI fresco. As culturas foram incubadas por 10 horas sob constante agitação a 37ºC. Medidas de densidade ótica a 600nm foram realizadas a cada hora. **(B)** Diluições seriadas das diferentes linhagens foram plaqueadas em BHI ágar para contagem de unidades formadoras de colônia (UFC) após 10 horas de incubação. Dados apresentados como média e desvio padrão de três diluições. A linhagem mutante Cp13 foi sempre crescida na presença do antibiótico canamicina (25 μg/mL).

5.2.2. Preparo do ligante de ferro 2,2-bipiridina (BIP) e determinação da concentração de uso

A concentração de uso do ligante bipiridina para utilização nos ensaios de depleção de ferro foi determinada utilizando meio BHI suplementado com com 200-220-240-260-280 μM da bipiridina (concentração final). O perfil de crescimento de culturas suplementadas com o agente quelante em relação ao meio BHI não suplementado foi utilizado para a determinar a condição de cultura de baixo ferro para as duas linhagens. De modo geral, o perfil de crescimento das duas linhagens foi similar, sendo possível observar uma redução no crescimento das duas linhagens em relação ao aumento da concentração do ligante bipiridina (Fig. 10). Dentre os valores de concentrações utilizados, os resultados de crescimento das concentrações de 240-260 μM se mostraram mais adequados, uma vez que em relação ao meio BHI-HI, essas concentrações reduziram a replicação bacteriana ao mesmo tempo que permitiram a produção suficiente de material para a extração de RNA (Fig. 10). Sendo assim, optou-se pela utilização de um valor intermediário (entre

240 e 260) de 250 µM do ligante bipiridina para os ensaios de restrição de ferro. Paralelamente, os dados demonstraram que diferenças no crescimento das linhagens em meio LI (suplementado com 240 e 260 µM de bipiridina) em relação ao meio HI começaram a ser notadas após 4 horas da inoculação, atingindo uma redução superior a 50-60% no valor da densidade celular de culturas HI para a linhagem Cp13 e 40-50% de redução para a linhagem T1 após 6 horas de incubação. Com base nesses resultados, optou-se por utilizar o início da fase exponencial (após 6h30min de incubação) para a extração do RNA das culturas LI e HI.

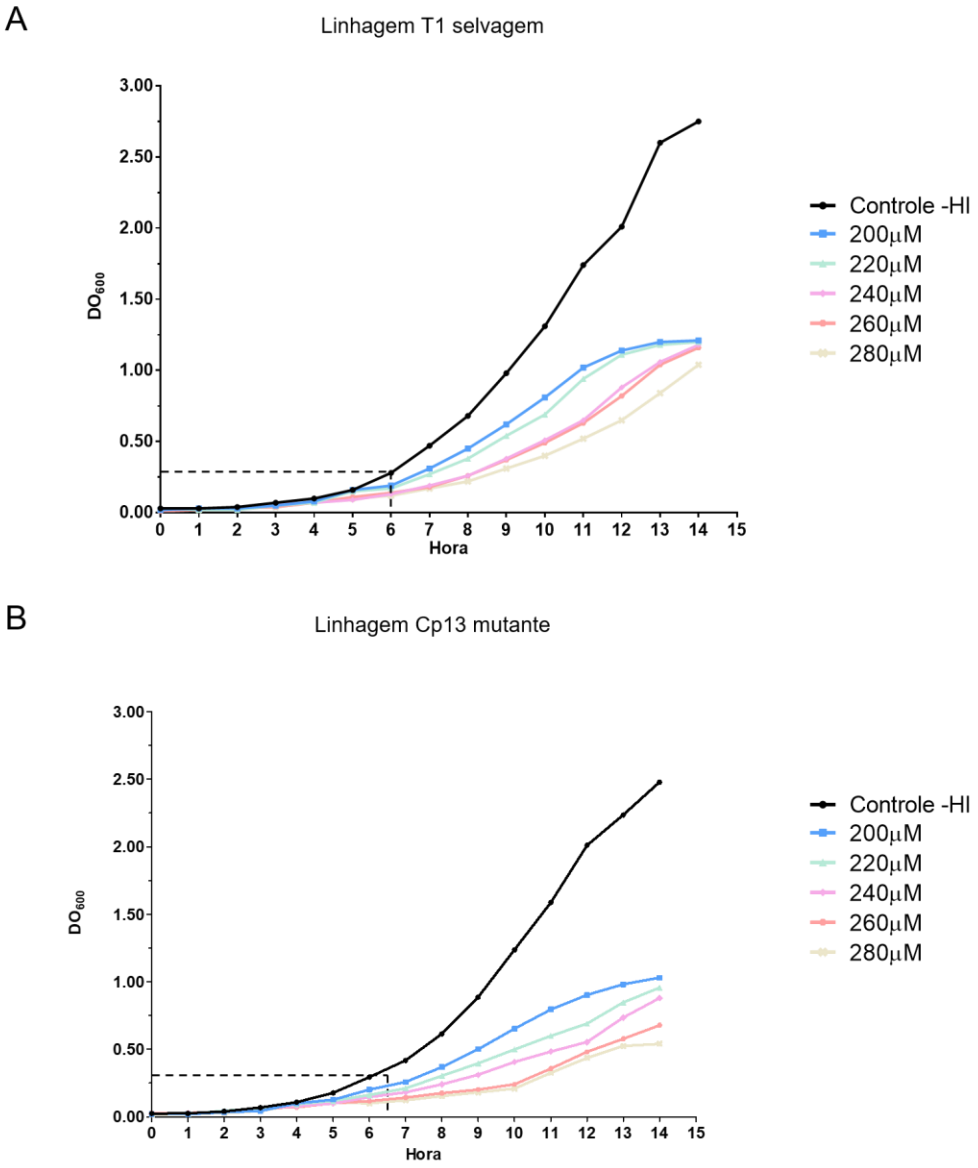


Figure 10. Curvas de crescimento das linhagens selvagem (T1) e mutante (Cp13) em meio BHI suplementado com diferentes concentrações do ligante 2,2-bipiridina

Culturas de *C. pseudotuberculosis* crescidas *overnight* em meio líquido foram diluídas a uma densidade ótica inicial de 0.02 em meio BHI fresco suplementado com diferentes concentrações do ligante 2,2-bipiridina (200, 220, 240, 260, 280 µM) e meio BHI controle (HI-controle). As culturas foram incubadas por 14 horas sob constante agitação a 37ºC. Medidas de densidade ótica a 600nm foram realizadas a cada hora. Curvas de crescimento da **(A)** linhagem selvagem T1 e **(B)** linhagem Cp13 mutante. A linhagem mutante Cp13 foi sempre crescida na presença do antibiótico canamicina (25 µg/mL).

5.2.3. Avaliação do crescimento das linhagens T1 parental e Cp13 mutante de C. pseudotuberculosis em meio (BHI) suplementado com 250 µM do ligante 2,2-bipiridina

A fim de avaliar o crescimento e viabilidade das linhagens de *C. pseudotuberculosis* à restrição de ferro, curvas de crescimento e contagem de unidades formadoras de colônia (UFC) foram realizadas com as linhagens selvagem (T1) e mutante (Cp13) em meio BHI com alta concentração de ferro, denominado meio HI, em relação ao meio BHI com baixa concentração de ferro, suplementada com 250 µM do agente quelante de ferro 2,2-bipiridina, denominado LI. A figura 5 mostra o crescimento bacteriano das linhagens, em relação a DO e a contagem por UFC/mL nos meios HI e LI. Como observado na figura 5, o crescimento bacteriano foi igualmente afetado quando as linhagens foram incubadas em meio LI, atingindo uma densidade celular final de 52% (T1) e 38.5% (CP13) do valor de densidade das culturas realizadas no meio HI, confirmando o estresse bacteriano em relação à restrição do metal (Fig. 11). De forma similar, uma redução significativa foi observada no número de colônias viáveis plaqueadas após incubação em meio LI para ambas as linhagens ($p < 0.05$). Em adição, culturas da linhagem mutante Cp13 de *C. pseudotuberculosis* crescidas em meio HI apresentaram taxa de crescimento semelhante a culturas da linhagem parental crescidas no mesmo meio (dados não apresentados).

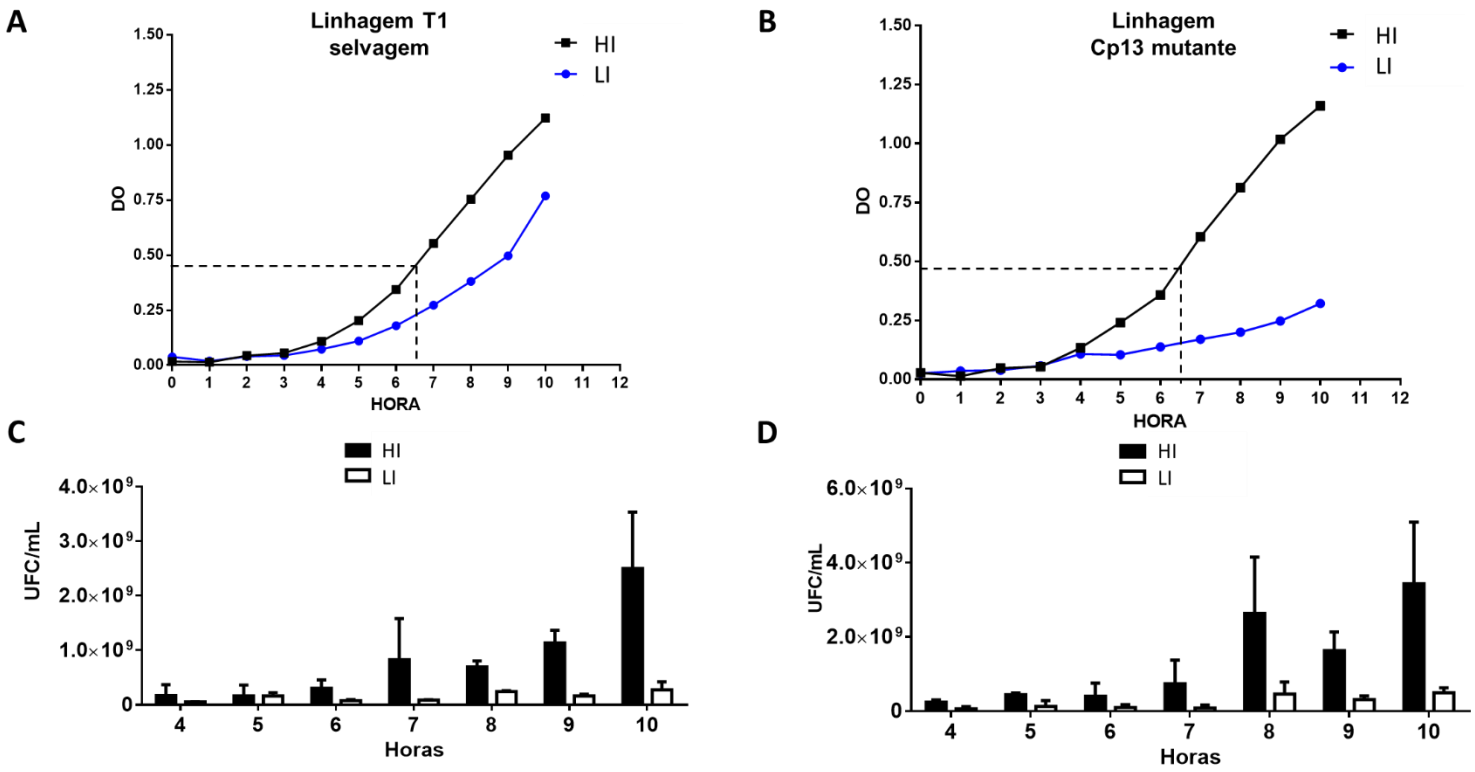


Figure 11. Curvas de crescimento das linhagens selvagem (T1) e mutante (Cp13) em meio BHI -HI e meio BHI suplementado com 250 µM do ligante 2,2-bipiridina.

(A/B) Culturas de *C. pseudotuberculosis* crescidas *overnight* em meio líquido foram diluídas a uma densidade ótica inicial de 0.02 em meio BHI fresco. As culturas foram incubadas por 10 horas sob constante agitação a 37ºC. Medidas de densidade ótica a 600nm foram realizadas a cada hora. **(C/D)** Diluições seriadas das diferentes linhagens foram plaqueadas em BHI ágar para contagem de unidades formadoras de colônia (UFC) a cada duas horas para avaliar a viabilidade bacteriana. Dados apresentados como média e desvio padrão de três diluições. Diferenças entre as condições experimentais foram consideradas significativas para $p < 0.05$. (Two-way ANOVA).

5.2.4. Análise da qualidade do RNA total

As amostras de RNA total foram coletadas após 6h30m da inoculação bacteriana nos meios. A quantificação e análise qualitativa do RNA total extraído foi avaliada pela razão Abs_{260}/Abs_{280} obtida utilizando espectrofotometria (NanoDrop, Thermo Scientific) e pelo perfil eletroforético das subunidades de rRNA obtido pelo equipamento Agilant 2100 Bioanalyzer. As amostras de RNA total extraídas apresentaram uma razão de absorbância (Abs_{260}/Abs_{280}) superior a 1.95, sendo consideradas adequadas para o sequenciamento. Em adição, as amostras de RNA total (AM1/2/3 da linhagem T1; AM2/3/4 da Cp13 mutante) também apresentaram pouco sinal de degradação, sendo possível visualizar a presença de duas bandas ribossomais íntegras (16S e 23S) (Fig. 12). As amostras HI e LI -AM1 da linhagem Cp13 foram processadas separadamente das demais amostras e apresentam, no gel, sinal de degradação com formação de fragmentos menores de RNA (Fig. 12). As amostras de RNA total foram submetidas a depleção do RNA ribossomal e preparação das bibliotecas para o sequenciamento.

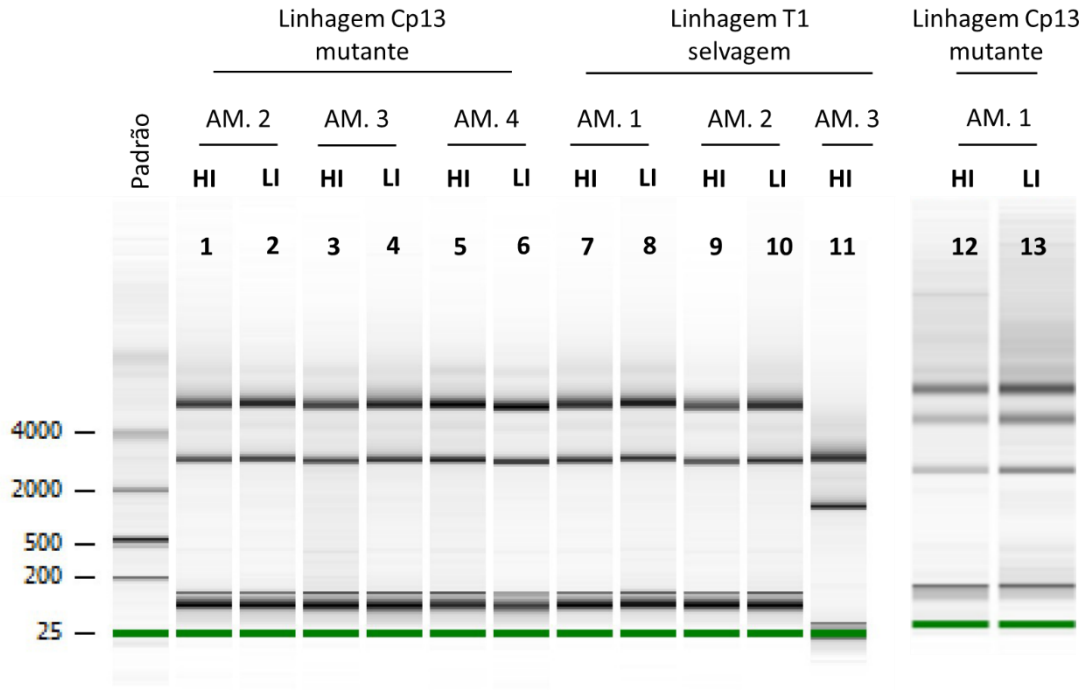


Figure 12. Análise qualitativa do perfil eletrofóretico dos RNAs totais extraídos de amostras da linhagem T1 selvagem e Cp13 mutante.

Gel dos RNAs das amostras HI e LI gerados pelo Agilent 2100 Bioanalyzer. Duas bandas claras são observadas no gel e representam a migração dos fragmentos dos rRNA16S e 23S. Degradação com formação de fragmentos menores de RNA podem ser observados nas amostras 12 e 13.

5.2.5. Estatísticas do sequenciamento e avaliação da qualidade das reads

Foram sequenciadas na plataforma Ion Torrent Proton System 4 réplicas biológicas de amostras pareadas (HI e LI) da linhagem Cp13 mutante e 3 réplicas biológicas de amostras pareadas da linhagem selvagem T1. Duas rodadas de sequenciamento foram realizadas na plataforma, gerando mais de 100 milhões de *single-end reads* com tamanho médio entre 8 - >300pb. Uma amostra HI da linhagem T1 apresentou um baixo rendimento total em relação ao sequenciamento, sendo sequenciadas 38.405 *reads* do par HI em comparação a 12.560.458 *reads* sequenciadas do par LI da mesma amostra. Como consequência, ambas amostras foram removidas do conjunto de resultados. Em relação a distribuição do conteúdo GC por *read*, foi possível identificar picos de conteúdo GC na distribuição real das sequências em relação à média global teórica de conteúdo GC do genoma, indicando a presença de sequências super-representadas nas bibliotecas sequenciadas. Contudo, a distribuição GC foi uniforme entre as bibliotecas, variando entre 53 e 55%.

De forma geral, as *reads* sequenciadas apresentaram um valor médio *phred* de qualidade por base entre 24-25. A pontuação *phred* é construída com base em uma série de preditores de qualidade, sendo utilizado para estimar a probabilidade de a base nomeada ter sido identificada de forma correta. Uma pontuação superior a 20 na escala *phred* é considerada aceitável para os dados gerados na plataforma Ion, e indica a probabilidade de ocorrência de 1 erro a cada 100 pb (99% de acurácia) (YUAN; XU; LEUNG, 2016). Em adição, foi observada uma redução significativa nos valores de qualidade por base na extremidade 3’ das *reads* (Fig. 13). As *reads* foram então processadas utilizando a opção -SLIDINGWINDOW do software Trimmomatic, resultando em uma perda de menos de 1% do valor total de *reads* em cada amostra. Essa opção minimiza a perda de dados brutos de alta qualidade ao cortar, na porção terminal 3' das *reads*, apenas quando a média da janela de leitura apresentar um valor de qualidade inferior ao *phred* selecionado (*phred* 10 – 90% acurácia) (BOLGER; LOHSE; USADEL, 2014). Ainda visando minimizar a perda de dados brutos, nenhum processamento foi realizado em relação ao tamanho das *reads* sequenciadas. Independentemente do tamanho das *reads*, a contagem do número de transcritos levaria em consideração apenas *reads* alinhadas de forma exclusiva a um gene do genoma de referência, eliminando assim alinhamentos espúrios causados por *reads* com sequências curtas. É importante ressaltar que apesar de ser uma etapa muito utilizada em analises transcricionais, o processamento de *reads* não é um processo ainda

bem definido, podendo variar muito entre os diversos trabalhos publicados (ANSELL *et al.*, 2013; BARRETT; DAVIS, 2012; STRAUB *et al.*, 2013; TAO *et al.*, 2013). Após o processamento, a qualidade das *reads* foi novamente avaliada levando em consideração a média de qualidade por base de cada sequência. De forma geral, as amostras processadas apresentaram um valor *phred* por sequência em torno de 25. A figura 8 apresenta uma visão geral dos dados após processamento. As *reads* processadas foram mapeados ao genoma de referência da linhagem 1002B de *C. pseudotuberculosis* utilizando um programa de alinhamento otimizado especificamente para dados produzidos pela plataforma Ion Torrent, o TMAP (Torrent Mapping Alignment Program). As bibliotecas sequenciadas tiveram um mapeamento geral de *reads* superior a 72% (Tabela 2). *Reads* com alinhamento ambíguo (múltiplo) foram excluídas da contagem. Dentre as *reads* mapeadas, 25-30% das *reads* alinhadas a genes exclusivos no genoma foram contadas e analisadas para expressão diferencial (Tabela 2). A baixa porcentagem de alinhamentos exclusivos já era esperada, uma vez que não foi possível otimizar o processo de depleção dos RNA ribossomais para essa bactéria.

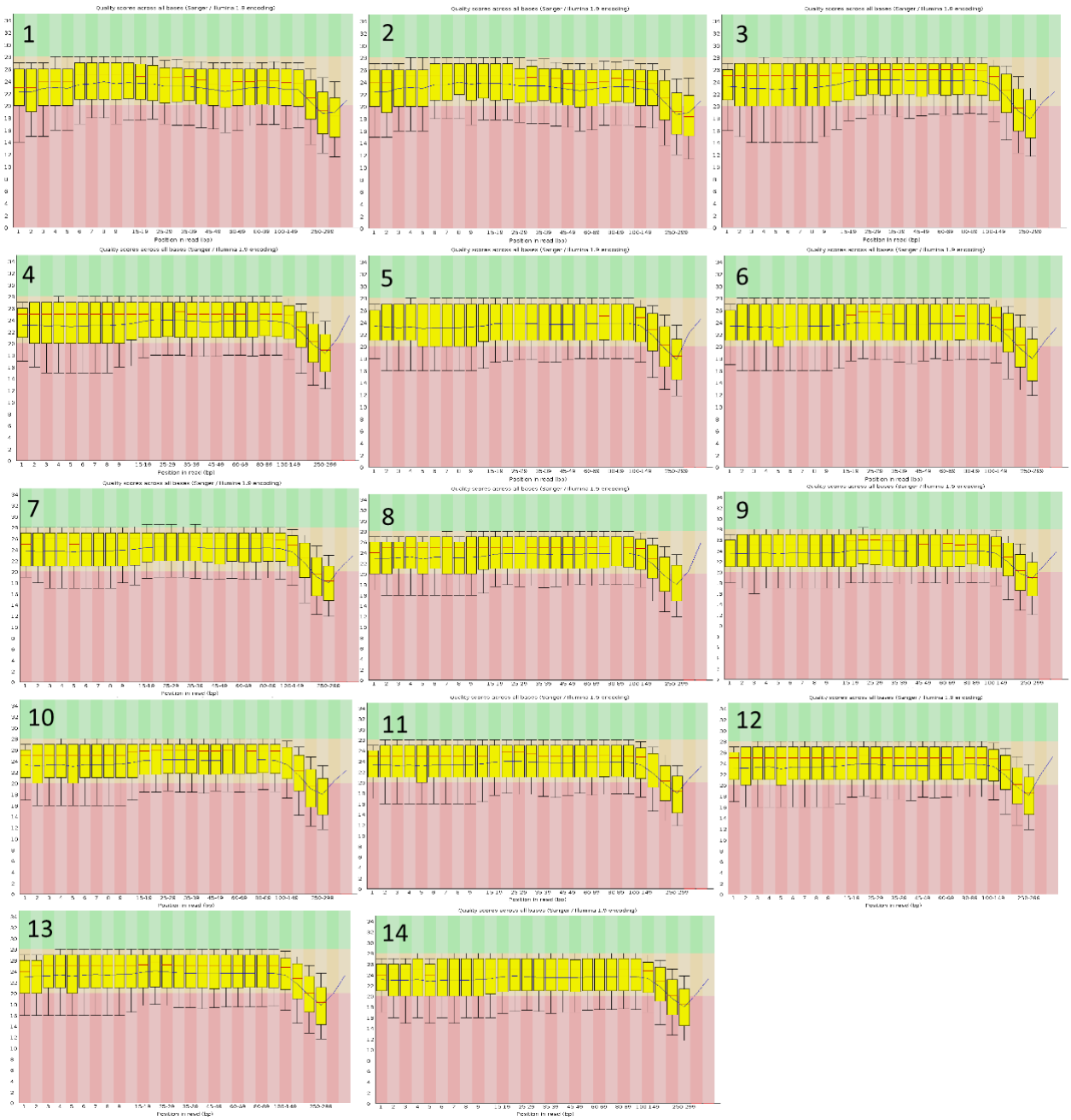


Figure 13. Valores de qualidade por base.

A linha vermelha central corresponde ao valor da mediana; as caixas amarelas representam o intervalo inter-quartil (25-75%); as barras superior e inferior representam os 10% e os 90% dos pontos; a linha azul representa a qualidade média das bases. Números de 1-8 representam as bibliotecas sequenciadas da linhagem Cp13 mutante; 9-14 bibliotecas sequenciadas da linhagem selvagem T1. Amostras HI: 1,3,5,7,9,11,13; Amostras LI: 2,4,6,8,10,12,14. Apenas uma amostra de cada replicata técnica foi apresentada.

Tabela 2: Estatística geral das bibliotecas sequenciadas de culturas HI (High Iron) e LI (Low Iron) das linhagens Cp13 mutante e T1 selvagem.

Tamanho médio das *reads* sequenciadas, média de conteúdo CG e número total de *reads* sequenciadas. Porcentagem de *reads* excluídas após aplicação do filtro de qualidade usando Trimmomatic. As *reads* processadas foram mapeadas ao genoma de referência utilizando o programa TMAP. *Reads* exclusivamente alinhadas a um único gene do genoma foram contadas pelo programa HTSeq-count.

	Id. amostras	Cond. Exp.	Tamanho (pb)	Média de conteúdo CG (%)	Batch	Total de *reads*	*Reads* processadas	Total de *reads* mapeadas	*Reads* mapeadas (%)	*Reads* não mapeadas	*Reads* com alinhamento multiplo (%)	*Reads* alinhadas exclu.	*Reads* exclusivas (%)
Cp13 mutante	S.1	HI	8-369	53	A	6014333	99.973	5934949	98.71	77775	71.29	1703889	28.7
	S.1	LI	8-375	53	A	10905772	99.970	10785269	98.92	117242	70.81	3147690	29.2
	S.2	HI_1	8-368	55	B	3936728	99.857	2997253	76.24	933853	72.67	819128	27.3
		HI _2	8-366	55	B	9035553	99.810	6720039	74.51	2298382	72.88	1822474	27.1
	S.2	LI	8-368	54	C	4834840	99.912	3505244	72.56	1325346	71.55	997153	28.4
	S.3	HI	8-374	55	C	8974704	99.885	7188306	80.19	1776069	71.68	2035506	28.3
	S.3	LI	8-378	55	C	4236781	99.923	3561273	84.12	672232	72.18	990776	27.8
	S.4	HI_1	8-368	54	C	1789770	99.921	1470238	82.21	318125	71.90	413162	28.1
		HI_2	8-369	54	C	3327609	99.916	2707354	81.43	617473	71.91	760389	28.1
	S.4	LI_1	8-369	55	C	2792977	99.963	2069063	74.11	722883	71.11	597748	28.9
		LI_2	8-364	55	C	6038287	99.957	4410477	73.07	1625205	71.12	1273762	28.9
T1 selvagem	S.1	HI_1	8-378	55	C	3988317	99.926	3414568	85.68	570796	71.32	979326	28.7
		HI_2	8-368	55	C	7886336	99.903	6655807	84.48	1222886	71.40	1903519	28.6
	S.1	LI_1	8-355	55	C	3537940	99.937	2624859	74.24	910842	71.18	756519	28.8
		LI_2	8-366	55	C	7801599	99.924	5682193	72.89	2113488	71.30	1630674	28.7
	S.2	HI	8-370	55	C	11975500	99.932	9624539	80.42	2342791	71.69	2724303	28.3
	S.2	LI	8-369	55	C	14575386	99.905	11897182	81.70	2664406	72.47	3275155	27.5
	S.3	HI	8-380	54	B	38385							
	S.3	LI	8-345	54	B	12560458							

5.2.6. Avaliação da qualidade dos dados e identificação do efeito de batch

A análise de componentes principais ou *principal component analysis* (PCA) foi utilizada para avaliar a relação entre as amostras. A função plotPCA do pacote R foi utilizada para representação gráfica da distribuição dos dados. Para a análise por componentes principais, uma matriz de análise, foi construída a partir dos dados de contagem normalizados das 12 amostras (HI e LI) de ambas as linhagens. Os dados foram transformados para escala logarítmica utilizando a função *rlog* (*regularized logarithm*) do pacote DESEq2. Efeito de batch foi identificado na análise dos componentes principais PC1 e PC2 (Fig. 14) e a medida estimada da dispersão para o agrupamento dos dados foi realizada utilizando a opção blind=false da função uma vez que já esperávamos um efeito de *batch* com grande variabilidade entre as contagens em função do processamento das amostras (LOVE, *et al*., 2014). Efeito de *batch* (batelada) é um subgrupo de medidas que apresentam variabilidade diferente entre as amostras e não possuem relação com a condição experimental biológica do estudo. Esse tipo de efeito é comum em dados de expressão gênica e pode ser, em grande parte, atribuído a diferenças técnicas experimentais causadas por alterações entre os dias de preparo e sequenciamento das amostras, lotes de reagentes e até mesmo diferenças entre equipamentos (LEEK et al., 2010). Variabilidade entre os *batches* amostrais representou 77% da variância total dos dados, e como esperado, reflete o agrupamento por batch entre as amostras 1 da linhagem Cp13 e as demais *batches* da mesma linhagem. Assim, acrescentamos a variável *batch* ao desenho experimental da análise de expressão da linhagem Cp13, de modo a controlar a contribuição experimental técnica na análise diferencial da linhagem. Apesar do efeito de *batch* ter sido significativo entre as amostras, no segundo componente observa-se a separação de grupos de amostras - HI e LI, evidenciando a separação entre as condições experimentais biológicas para as duas linhagens (8% da variância total). Na figura 14, os valores PC1 e PC2 representam 85% da variância total dos dados. O agrupamento por *batch* na linhagem Cp13 também foi confirmado pela distância Euclidiana calculada com os dados brutos de contagem normalizados e transformados pele função *rlog* no R (Fig. 14B)

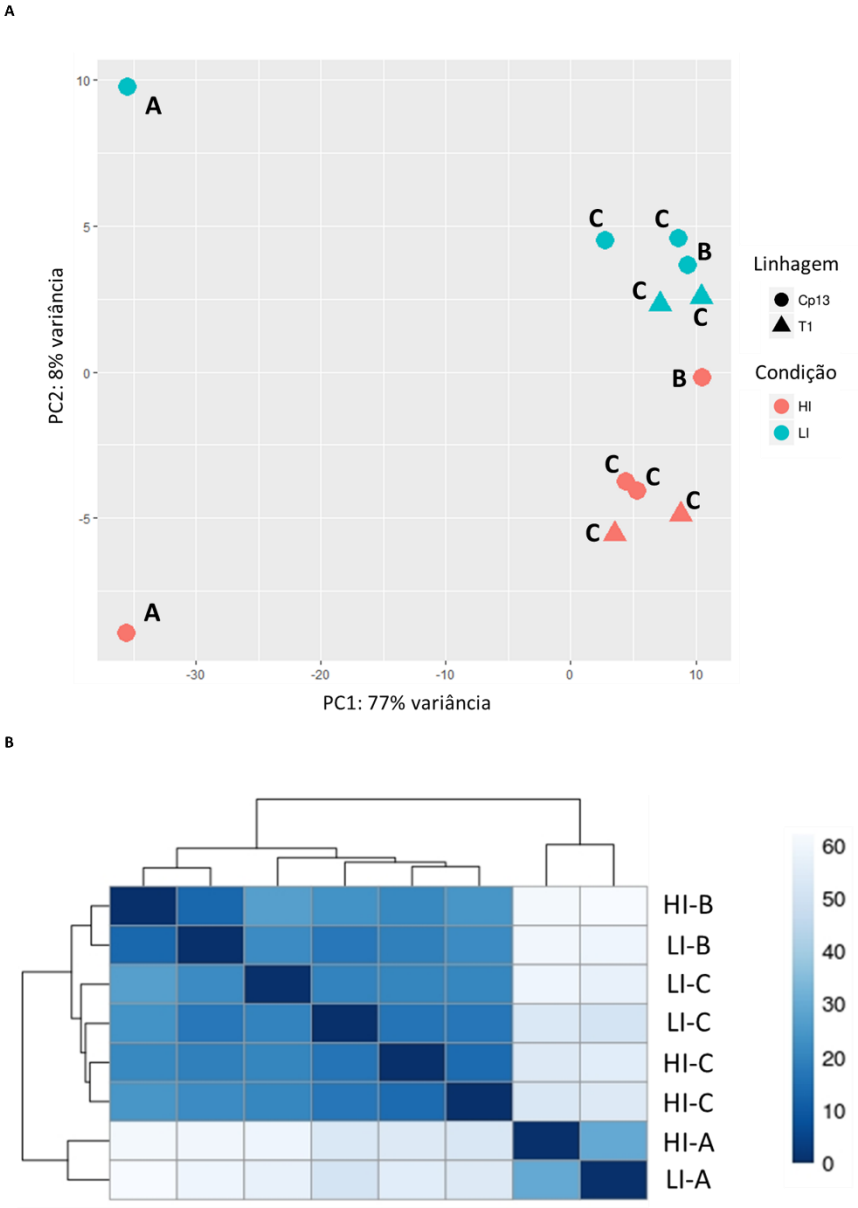


Figure 14. Análise por componentes principais (PCA) das amostras HI e LI da linhagem Cp13 mutante e T1 selvagem e distância Euclidiana das amostras da linhagem Cp13 mutante

Dados brutos de contagem normalizados transformados para escala logarítmica (*rlog*) e analisados por PCA e distância euclidiana. **(A)** Componentes PC1 e PC2 são apresentados e representam 85% da variância total do sistema. As linhagens foram representadas por figuras geométricas: (▲) T1 selvagem e (●) Cp13 mutante. *Batches* foram identificados de A-C. As condições experimentais HI e LI foram representadas por cores: HI -:laranja; LI - azul. O efeito de *batch* foi atribuído ao preparo das amostras e a variável *batch* foi incluída no desenho experimental da ferramenta DESeq2. **(B)** Heatmap da distância Euclidiana entre as amostras da linhagem Cp13 mutante demonstrando a organização hierárquica das amostras.

5. CONCLUSÕES

Os resultados obtidos nesse estudo corroboram com a hipótese de que a percepção bacteriana à disponibilidade de ferro desempenha um papel essencial na regulação transcricional de genes associados a homeostase de ferro, bem como a virulência de *Corynebacterium pseudotuberculosis*. Em resumo, as análises de expressão gênica da linhagem tipo Selvagem (T1) e da linhagem mutante para o gene *ciuA* (linhagem Cp13) de *C. pseudotuberculosis* contribuíram para uma melhor compreensão dos genes cuja expressão é regulada pela disponibilidade de ferro. Dentre os GDEs identificados compartilhados entre as duas linhagens, foi possível observar a importância da captação de ferro na forma de heme e o possível envolvimento desses sistemas na aquisição de ferro e virulência de *C. pseudotuberculosis*. Notavelmente, nossos resultados indicam que a restrição de ferro compromete o metabolismo aeróbico bacteriano e sugerem que o controle da demanda intracelular de ferro é essencial na resposta adaptativa de *C. pseudotuberculosis*. De forma geral, o perfil transcricional da linhagem T1 selvagem apontou que a restrição de ferro comprometeu o nível de atividade metabólica da linhagem, conducente à restrição nutricional do ambiente. Comparativamente, a interrupção do gene *ciuA* na linhagem mutante levou a uma resposta gênica centrada na aquisição do metal, sinalizando um mecanismo de compensação entre os sistemas de captação de ferro de *C. pseudotuberculosis*. Esse resultado pode indicar que a bactéria não restringe a captação de ferro a uma única estratégia, mas pode ajustar a aquisição de ferro de acordo com a demanda celular. Infelizmente, nossos resultados não possibilitaram a avaliação da regulação de sideróforos em resposta à restrição de ferro; contudo, a aquisição de ferro mediada por sideróforos parece contribuir para o crescimento bacteriano. Concludentemente, fica claro que a importância da proteína CiuA dentre os sistemas de homeostase de ferro em *C. pseudotuberculosis*, bem como o papel da associação entre o gene *ciuA* e os demais genes identificados DEs na linhagem mutante precisa ainda ser explorado. Certamente, a complexidade das interações entre a bactéria e o hospedeiro não podem ser ignorados, abrindo perspectiva para análises futuras envolvendo essa linhagem mutante. Conclusivamente, nossos resultados do perfil transcricional juntamente com a rede de regulação aqui proposta, indicam que o processo de resposta bacteriana à disponibilidade de ferro envolve complexas interações entre genes regulados por ferro e fatores transcricionais, de modo a permitir uma rápida adaptação bacteriana as condições ambientais. Nossos resultados fornecem uma análise abrangente sobre a resposta transcricional de *C. pseudotuberculosis* sob restrição de ferro,

contribuindo assim para o desenvolvimento de estudos futuros voltados para a identificação de novos alvos terapêuticos contra doenças causadas por este importante microrganismo.

6. PERSPECTIVAS

Os resultados obtidos neste trabalho permitiram uma avaliação abrangente da resposta de *C. pseudotuberculosis* em relação à restrição de ferro. No entanto, nossos dados em relação a regulação de sideróforos não foram suficientes para esclarecer o envolvimento desse sistema na aquisição de ferro em *C. pseudotuberculosis*. Em adição, os resultados obtidos do perfil transcricional da linhagem mutante não nos permitiram esclarecer o papel da proteína CiuA e sua associação com os resultados dos ensaios vacinais obtidos com a linhagem mutante, bem como seu envolvimento com os diversos outros sistemas de aquisição de ferro em *C. pseudotuberculosis*. Sendo assim esse trabalho abre perspectivas para realizar a quantificação da expressão dos genes *ciu* e *fag* de *C. pseudotuberculosis* por PCR em tempo real, expondo a bactéria a proteínas extracelulares de ligação ao ferro do hospedeiro como a transferrina e lactoferrina, tornando assim possível avaliar a contribuição desses sistemas na aquisição de ferro. Em adição, esse experimento poderia contribuir significativa para avaliar a hipótese envolvendo a dependência de proteínas do hospedeiro na indução transcricional de genes do sistema de aquisição por sideróforos.

Apesar da importância de doenças causadas por *C. pseudotuberculosis* na agricultura mundial, não há diagnostico subclínica ou vacinas que sejam realmente eficientes no combate a LC. Desse modo, os dados aqui apresentados abrem perspectivas para a avaliação da capacidade imunogênica de algumas das proteínas secretadas envolvidas na aquisição de ferro heme como alvo potencial para o desenvolvimento vacinal ou imunobiológico para o controle da LC. O interesse por esse subgrupo de produtos secretados é devido ao fato desses produtos gênicos constituírem a principal classe de fatores de virulência bacterianos, bem como por terem apresentado os valores mais elevados de indução entre os GDEs em ambas as linhagens. Análises preliminares serão realizadas para avaliar o nível de conservação dessas proteínas dentre as linhagens de *C. pseudotuberculosis* e a caracterização do potencial imunogênico desses produtos

proteicos serão avaliados *in silico* (predição de epítopos). O potencial protetor dessas proteínas será avaliado em modelos de infecção em camundongos

7. REFERÊNCIAS BIBLIOGÁFICAS

AGUIAR, André *et al.* Mecanismo e aplicações da reação de fenton assistida por compostos fenólicos redutores de ferro. *Química Nova*, v. 30, n. 3, p. 623–628, jun. 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422007000300023&lng=pt&nrm=iso&tlng=pt>.

ALAMURI, Praveen; MAIER, Robert J. Methionine sulphoxide reductase is an important antioxidant enzyme in the gastric pathogen Helicobacter pylori. *Molecular Microbiology*, v. 53, n. 5, p. 1397–1406, set. 2004.

ALEMAN, M *et al.* Corynebacterium pseudotuberculosis infection in horses: 538 cases (1982-1993). *Journal of the American Veterinary Medical Association*, v. 209, n. 4, p. 804–9, 15 ago. 1996. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8756884>.

ALVES, F. S. F.; SANTIAGO, L. B.; PINHEIRO, R. R. *Linfadenite caseosa: o estado da arte.* . [S.l: s.n.], 2007.

ANDREINI, Claudia *et al.* Metal ions in biological catalysis: from enzyme databases to general principles. *JBIC Journal of Biological Inorganic Chemistry*, v. 13, n. 8, p. 1205–1218, 5 nov. 2008. Disponível em: <http://link.springer.com/10.1007/s00775-008-0404-5>.

ANDREWS, Simon C; ROBINSON, Andrea K; RODRÍGUEZ-QUIÑONES, Francisco. Bacterial iron homeostasis. *FEMS microbiology reviews*, v. 27, n. 2–3, p. 215–237, 2003. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12829269>.

ANSELL, Brendan R.E. *et al.* Insights into the immuno-molecular biology of Angiostrongylus vasorum through transcriptomics—Prospects for new interventions. *Biotechnology Advances*, v. 31, n. 8, p. 1486–1500, dez. 2013. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0734975013001201>.

AQUINO DE SÁ, Maria da Conceição *et al.* Distribution of PLD and FagA, B, C and D genes in Corynebacterium pseudotuberculosis isolates from sheep and goats with caseus lymphadenitis. *Genetics and Molecular Biology*, v. 36, n. 2, p. 265–268, jul. 2013.

ARSENAULT, Julie *et al.* Prevalence of and carcass condemnation from maedi-visna, paratuberculosis and caseous lymphadenitis in culled sheep from Quebec, Canada. *Preventive veterinary medicine*, v. 59, n. 1–2, p. 67–81, 30 maio 2003. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12719018>.

ASANO, R. L.; DAVIES, J. Molecular characterization of the thioredoxin system of Mycobacterium smegmatis. *Research in Microbiology*, v. 149, n. 8, p. 567–576, set. 1998.

AZIZ, Ramy K. *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, v. 9, p. 75, 2008.

BAIRD, G. J.; FONTAINE, M. C. Corynebacterium pseudotuberculosis and its Role in Ovine Caseous Lymphadenitis. *Journal of Comparative Pathology*, v. 137, n. 4, p. 179–210, 2007a.

BAIRD, G.J.; FONTAINE, M.C. Corynebacterium pseudotuberculosis and its Role in Ovine Caseous Lymphadenitis. *Journal of Comparative Pathology*, v. 137, n. 4, p. 179–210, nov. 2007b. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0021997507001211>.

BANDYOPADHYAY, P.; STEINMAN, H. M. Catalase-peroxidases of Legionella pneumophila: cloning of the katA gene and studies of KatA function. *Journal of Bacteriology*, v. 182, n. 23, p. 6679–6686, dez. 2000.

BARRETT, Craig F.; DAVIS, Jerrold I. The plastid genome of the mycoheterotrophic Corallorhiza striata (Orchidaceae) is in the relatively early stages of degradation. *American Journal of Botany*, v. 99, n. 9, p. 1513–1523, set. 2012. Disponível em: <http://doi.wiley.com/10.3732/ajb.1200256>.

BECKER, Kyle W; SKAAR, Eric P. Metal limitation and toxicity at the interface between host and pathogen. *FEMS microbiology reviews*, v. 38, n. 6, p. 1235–49, nov. 2014. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25211180>.

BELCHIOR, Silvia Estevao *et al.* ACTUALIZACIÓN SOBRE LINFOADENITIS CASEOSA: EL AGENTE ETIOLÓGICO Y LA ENFERMEDAD. *Veterinaria Argentina*, 2006.

BIERNE, Hélène *et al.* Inactivation of the srtA gene in Listeria monocytogenes inhibits anchoring of surface proteins and affects virulence. *Molecular Microbiology*, v. 43, n. 4, p. 869–881, fev. 2002.

BILLINGTON, Stephen J.; ESMAY, Paula A.; SONGER, J.; JOST, B.; *et al.* Identification and role in virulence of putative iron acquisition genes from Corynebacterium pseudotuberculosis. *FEMS Microbiology Letters*, v. 208, n. 1, p. 41–45, 2002.

BILLINGTON, Stephen J.; ESMAY, Paula A.; SONGER, J. Glenn; JOST, B. Helen. Identification and role in virulence of putative iron acquisition genes from Corynebacterium pseudotuberculosis. *FEMS microbiology letters*, v. 208, n. 1, p. 41–45, fev. 2002.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, v. 30, n. 15, p. 2114–2120, 1 ago. 2014. Disponível em: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu170>.

BRAIBANT, Martine; GUILLOTEAU, Laurence; ZYGMUNT, Michel S. Functional characterization of Brucella melitensis NorMI, an efflux pump belonging to the multidrug and toxic compound extrusion family. *Antimicrobial Agents and Chemotherapy*, v. 46, n. 9, p. 3050–3053, set. 2002.

BRAUN, V. Iron uptake mechanisms and their regulation in pathogenic bacteria. *International journal of medical microbiology : IJMM*, v. 291, n. 2, p. 67–79, 2001. Disponível em: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11437341&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1078/1438-4221-00103>.

BROWN, Jeremy S; HOLDEN, David W. Iron acquisition by Gram-positive bacterial pathogens. *Microbes and infection*, v. 4, n. 11, p. 1149–56, set. 2002. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12361915>.

BRUNE, Iris *et al.* The DtxR protein acting as dual transcriptional regulator directs a global regulatory network involved in iron metabolism of Corynebacterium glutamicum. *BMC genomics*, v. 7, p. 21, 9 fev. 2006. Disponível em:

<http://www.ncbi.nlm.nih.gov/pubmed/16469103>.

BULLEN, J.J. E GRIFFITHS, E. Iron Binding Proteins and Host Defense. *Iron and Infection*. 2. ed. New York: John Wiley and Sons Ltd, 1999. p. 327–368.

BURKOVSKI, Andreas. Cell Envelope of Corynebacteria: Structure and Influence on Pathogenicity. *ISRN Microbiology*, v. 2013, p. 1–11, 2013. Disponível em: <https://www.hindawi.com/archive/2013/935736/>.

CALHOUN, L. N.; KWON, Y. M. Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in Escherichia coli: a review. *Journal of Applied Microbiology*, v. 110, n. 2, p. 375–386, fev. 2011.

CAMERON, CM; SWART, CF. A new liquid medium for the cultivation of Corynebacterium pseudotuberculosis. *J S Afr Vet Med Ass*, v. 36, p. 185–188, 1965.

CARPENTER, B. M.; WHITMIRE, J. M.; MERRELL, D. S. This Is Not Your Mother's Repressor: the Complex Role of Fur in Pathogenesis. *Infection and Immunity*, v. 77, n. 7, p. 2590–2601, 1 jul. 2009. Disponível em: <http://iai.asm.org/cgi/doi/10.1128/IAI.00116-09>.

CARVALHO, Daiane M. *et al.* Reference genes for RT-qPCR studies in Corynebacterium pseudotuberculosis identified through analysis of RNA-seq data. *Antonie van Leeuwenhoek*, v. 106, n. 4, p. 605–614, out. 2014.

CARVER, Tim J. *et al.* ACT: the Artemis Comparison Tool. *Bioinformatics (Oxford, England)*, v. 21, n. 16, p. 3422–3423, ago. 2005.

CASONATO, Stefano *et al.* Mycobacterium tuberculosis requires the ECF sigma factor SigE to arrest phagosome maturation. *PloS One*, v. 9, n. 9, p. e108893, 2014.

CASSAT, James E.; SKAAR, Eric P. Iron in Infection and Immunity. *Cell Host & Microbe*, v. 13, n. 5, p. 509–519, maio 2013. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1931312813001522>.

CHASTANET, Arnaud *et al.* clpB, a novel member of the Listeria monocytogenes CtsR regulon, is involved in virulence but not in general stress tolerance. *Journal of Bacteriology*, v. 186, n. 4, p. 1165–1174, fev. 2004.

CHEVREUX, B.; WETTER, T.; SUHAI, S. *Genome Sequence Assembly Using Trace Signals and Additional Sequence Information*. [S.l: s.n.], [S.d.].

CHOI, S. H.; BAUMLER, D. J.; KASPAR, C. W. Contribution of dps to acid stress tolerance and oxidative stress tolerance in Escherichia coli O157:H7. *Applied and Environmental Microbiology*, v. 66, n. 9, p. 3911–3916, set. 2000.

COELHO, Keila da Silva. *Isolamento, clonagem e caracterização molecular do gene hsp60 de Corynebacterium pseudotuberculosis e sua utilização na construção de uma vacina de de DNA e de subunidade protéica.* 2007. 120 f. Universidade Federal de Minas Gerais, 2007.

COSTA, CARLOTA. Comprehensive molecular screening: from the RT-PCR to the RNA-seq. v. 2(2), p. 87–91, [S.d.].

COSTA, L.F.M. Corynebacterium pseudotuberculosis, o agente etiológico da linfadenite caseosa em caprinos. *Rev Cienc Med Biol*, v. 1, n. 1, p. 105–115, 2010.

COSTERTON, J. W.; STEWART, P. S.; GREENBERG, E. P. Bacterial biofilms: a common cause of persistent infections. *Science (New York, N.Y.)*, v. 284, n. 5418, p. 1318–1322, maio 1999.

CRAIG, Jane E.; NOBBS, Angela; HIGH, Nicola J. The extracytoplasmic sigma factor, final sigma(E), is required for intracellular survival of nontypeable Haemophilus influenzae in J774 macrophages. *Infection and Immunity*, v. 70, n. 2, p. 708–715, fev. 2002.

D'AQUINO, J. A. *et al.* Mechanism of metal ion activation of the diphtheria toxin repressor DtxR. *Proceedings of the National Academy of Sciences*, v. 102, n. 51, p. 18408–18413, 20 dez. 2005. Disponível em: <http://www.pnas.org/cgi/doi/10.1073/pnas.0500908102>.

DERVEAUX, Stefaan; VANDESOMPELE, Jo; HELLEMANS, Jan. How to do successful gene expression analysis using real-time PCR. *Methods*, v. 50, n. 4, p. 227–230, abr. 2010.

DOHERR, M G *et al.* Risk factors associated with Corynebacterium pseudotuberculosis infection in California horses. *Preventive veterinary medicine*, v. 35, n. 4, p. 229–39, 30 jun. 1998. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9689656>.

DOIG, Peter *et al.* Rational design of inhibitors of the bacterial cell wall synthetic enzyme GlmU using virtual screening and lead-hopping. *Bioorganic & Medicinal Chemistry*, v. 22, n. 21, p. 6256–6269, nov. 2014.

DORELLA, F. A. *et al.* In Vivo Insertional Mutagenesis in Corynebacterium pseudotuberculosis: an Efficient Means To Identify DNA Sequences Encoding Exported Proteins. *Applied and Environmental Microbiology*, v. 72, n. 11, p. 7368–7372, 1 nov. 2006. Disponível em: <http://aem.asm.org/cgi/doi/10.1128/AEM.00294-06>.

DORELLA, Fernanda A *et al.* Antigens of Corynebacterium pseudotuberculosis and prospects for vaccine development. *Expert Review of Vaccines*, v. 8, n. 2, p. 205–213, 9 fev. 2009. Disponível em: <http://www.tandfonline.com/doi/full/10.1586/14760584.8.2.205>.

DORELLA, Fernanda Alves *et al.* Corynebacterium pseudotuberculosis : microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Veterinary Research*, v. 37, n. 2, p. 201–218, mar. 2006. Disponível em: <http://www.edpsciences.org/10.1051/vetres:2005056>.

DOUGLAS, T. *et al.* Methionine sulfoxide reductase A (MsrA) deficiency affects the survival of Mycobacterium smegmatis within macrophages. *Journal of Bacteriology*, v. 186, n. 11, p. 3590–3598, jun. 2004.

FACCIOLI-MARTINS, P.Y.; ALVES, F.S.F.; PINHEIRO, R.R. *Linfadenite caseosa: perspectivas no diagnóstico, tratamento e controle*. . [S.l: s.n.], 2014.

FARR, S. B.; KOGOMA, T. Oxidative stress responses in Escherichia coli and Salmonella typhimurium. *Microbiological Reviews*, v. 55, n. 4, p. 561–585, dez. 1991.

FERREIRA, A.; O'BYRNE, C. P.; BOOR, K. J. Role of sigma(B) in heat, ethanol, acid, and oxidative stress resistance and during carbon starvation in Listeria monocytogenes. *Applied and Environmental Microbiology*, v. 67, n. 10, p. 4454–4457, out. 2001.

FOLLMANN, Martin *et al.* Functional genomics of pH homeostasis in Corynebacterium glutamicum revealed novel links between pH response, oxidative stress, iron homeostasis and methionine synthesis. *BMC genomics*, v. 10, p. 621, 2009.

FUKUSHIMA, Tatsuya; ALLRED, Benjamin E.; RAYMOND, Kenneth N. Direct evidence of iron uptake by the Gram-positive siderophore-shuttle mechanism without iron reduction. *ACS Chemical Biology*, v. 9, n. 9, p. 2092–2100, 2014.

GAL-MOR, Ohad; FINLAY, B. Brett. Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cellular Microbiology*, v. 8, n. 11, p. 1707–1719, nov. 2006.

GALARDINI, Marco *et al.* CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code for Biology and Medicine*, v. 6, p. 11, 2011.

GIBSON, C. M.; CAPARON, M. G. Alkaline Phosphatase Reporter Transposon for Identification of Genes Encoding Secreted Proteins in Gram-Positive Microorganisms. *Applied and Environmental Microbiology*, v. 68, n. 2, p. 928–932, 1 fev. 2002. Disponível em: <http://aem.asm.org/cgi/doi/10.1128/AEM.68.02.928-932.2002>.

GIORGIO, Marco *et al.* Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nature Reviews. Molecular Cell Biology*, v. 8, n. 9, p. 722–728, set. 2007.

GOMES, Débora L. R. *et al.* Corynebacterium diphtheriae as an emerging pathogen in nephrostomy catheter-related infection: evaluation of traits associated with bacterial virulence. *Journal of Medical Microbiology*, v. 58, n. Pt 11, p. 1419–1427, nov. 2009.

GUIMARÃES, Alessandro de Sá *et al.* Molecular characterization of Corynebacterium pseudotuberculosis isolates using ERIC-PCR. *Veterinary Microbiology*, v. 153, n. 3–4, p. 299–306, dez. 2011.

HAAS, Brian J. *et al.* How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC genomics*, v. 13, p. 734, 2012.

HARD, G C. Comparative toxic effect of the surface lipid of Corynebacterium ovis on peritoneal macrophages. *Infection and immunity*, v. 12, n. 6, p. 1439–49, dez. 1975. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/1205621>.

HARE, Rebekah F.; HUEFFER, Karsten. Francisella novicida pathogenicity island encoded proteins were secreted during infection of macrophage-like cells. *PloS One*, v. 9, n. 8, p. e105773, 2014.

HARTMANN, A; BRAUN, V. Iron uptake and iron limited growth of Escherichia coli K-12. *Archives of microbiology*, v. 130, n. 5, p. 353–6, 2 dez. 1981. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7034667>.

HODGSON, A L *et al.* Rational attenuation of Corynebacterium pseudotuberculosis: potential cheesy gland vaccine and live delivery vehicle. *Infection and immunity*, v. 60, n. 7, p. 2900–5, jul. 1992. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/1612756>.

HONARMAND EBRAHIMI, Kourosh; HAGEDOORN, Peter-Leon; HAGEN, Wilfred R. Unity in the Biochemistry of the Iron-Storage Proteins Ferritin and Bacterioferritin. *Chemical Reviews*, v. 115, n. 1, p. 295–326, 14 jan. 2015. Disponível em: <http://pubs.acs.org/doi/10.1021/cr5004908>.

HOOD, M. Indriati; SKAAR, Eric P. Nutritional immunity: Transition metals at the pathogen-host interface. *Nature Reviews Microbiology*, v. 10, n. 8, p. 525–537, 2012. Disponível em: <http://dx.doi.org/10.1038/nrmicro2836>.

ISHIKAWA, Morio *et al.* Cloning and characterization of clpB in Acetobacter pasteurianus NBRC 3283. *Journal of Bioscience and Bioengineering*, v. 110, n. 1, p. 69–71, jul. 2010.

JANKUTE, Monika *et al.* Assembly of the Mycobacterial Cell Wall. *Annual review of microbiology*, v. 69, p. 405–23, 2015. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/26488279>.

JOLLY, R. D. Some Observations on Surface Lipids of Virulent and Attenuated Strains of Corynebacterium ovis. *Journal of Applied Bacteriology*, v. 29, n. 1, p. 189–196, abr. 1966. Disponível em: <http://doi.wiley.com/10.1111/j.1365-2672.1966.tb03467.x>.

JONES, T.C.; HUNT, R.D.; KING, N.W. Moléstias Causadas por Bactérias. *Patologia Veterinária*. 6. ed. São Paulo: Manolo, 2000. p. 489–491.

JOZEFCZUK, Szymon *et al.* Metabolomic and transcriptomic stress response of Escherichia coli. *Molecular Systems Biology*, v. 6, p. 364, maio 2010.

KAES, Christian; KATZ, Alexander; HOSSEINI, Mir Wais. Bipyridine: The Most Widely Used Ligand. A Review of Molecules Comprising at Least Two 2,2‘-Bipyridine Units. *Chemical Reviews*, v. 100, n. 10, p. 3553–3590, out. 2000. Disponível em: <http://pubs.acs.org/doi/abs/10.1021/cr990376z>.

KALINOWSKI, D. S. The Evolution of Iron Chelators for the Treatment of Iron Overload Disease and Cancer. *Pharmacological Reviews*, v. 57, n. 4, p. 547–583, 1 dez. 2005. Disponível em: <http://pharmrev.aspetjournals.org/cgi/doi/10.1124/pr.57.4.2>.

KANNAN, T. R. *et al.* Characterization of a unique ClpB protein of Mycoplasma pneumoniae and its impact on growth. *Infection and Immunity*, v. 76, n. 11, p. 5082–5092, nov. 2008.

KARIUKI, D. P.; POULTON, J. Corynebacterial infection of cattle in Kenya. *Tropical Animal Health and Production*, v. 14, n. 1, p. 33–36, fev. 1982. Disponível em: <http://link.springer.com/10.1007/BF02281102>.

KAZMIERCZAK, Mark J.; WIEDMANN, Martin; BOOR, Kathryn J. Alternative sigma factors and their roles in bacterial virulence. *Microbiology and molecular biology reviews: MMBR*, v. 69, n. 4, p. 527–543, dez. 2005.

KIM, J. K. *et al.* Bacterial Cell Wall Synthesis Gene uppP Is Required for Burkholderia Colonization of the Stinkbug Gut. *Applied and Environmental Microbiology*, v. 79, n. 16, p. 4879–4886, ago. 2013.

KUNKLE, C. A.; SCHMITT, M. P. Analysis of a DtxR-Regulated Iron Transport and Siderophore Biosynthesis Gene Cluster in Corynebacterium diphtheriae. *Journal of Bacteriology*, v. 187, n. 2, p. 422–433, 15 jan. 2005. Disponível em: <http://jb.asm.org/cgi/doi/10.1128/JB.187.2.422-433.2005>.

KUNKLE, C. A.; SCHMITT, M. P. Analysis of the Corynebacterium diphtheriae DtxR Regulon: Identification of a Putative Siderophore Synthesis and Transport System That Is Similar to the Yersinia High-Pathogenicity Island-Encoded Yersiniabactin Synthesis and Uptake System. *Journal of Bacteriology*, v. 185, n. 23, p. 6826–6840, 1 dez. 2003.

Disponível em: <http://jb.asm.org/cgi/doi/10.1128/JB.185.23.6826-6840.2003>.

LADETTO, M. *et al.* Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia*, v. 28, n. 6, p. 1299–1307, jun. 2014.

LEE, Ji-Sun *et al.* KatA, the major catalase, is critical for osmoprotection and virulence in Pseudomonas aeruginosa PA14. *Infection and Immunity*, v. 73, n. 7, p. 4399–4403, jul. 2005.

LIU, He *et al.* Regulatory role of tetR gene in a novel gene cluster of Acidovorax avenae subsp. avenae RS-1 under oxidative stress. *Frontiers in Microbiology*, v. 5, p. 547, 2014.

LOURDAULT, Kristel *et al.* Inactivation of clpB in the pathogen Leptospira interrogans reduces virulence and resistance to stress conditions. *Infection and Immunity*, v. 79, n. 9, p. 3711–3717, set. 2011.

LUIZ DE PAULA CASTRO, Thiago. *Avaliação da expressão diferencial de genes codificadores de fatores sigma de corynebacterium pseudotuberculosis em resposta a agentes geradores de estresse oxidativo.* 2009. 119 f. Universidade Federal de Minas Gerais, 2009.

MANGANELLI, R. *et al.* The Mycobacterium tuberculosis ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. *Molecular Microbiology*, v. 41, n. 2, p. 423–437, jul. 2001.

MARESSO, Anthony W.; GARUFI, Gabriella; SCHNEEWIND, Olaf. Bacillus anthracis Secretes Proteins That Mediate Heme Acquisition from Hemoglobin. *PLoS Pathogens*, v. 4, n. 8, p. e1000132, 22 ago. 2008. Disponível em: <http://dx.plos.org/10.1371/journal.ppat.1000132>.

MAZMANIAN, Sarkis K *et al.* Passage of heme-iron across the envelope of Staphylococcus aureus. *Science (New York, N.Y.)*, v. 299, n. 5608, p. 906–9, 7 fev. 2003. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12574635>.

MCCARTHY, Ronan R. *et al.* A new regulator of pathogenicity (bvlR) is required for full virulence and tight microcolony formation in Pseudomonas aeruginosa. *Microbiology (Reading, England)*, v. 160, n. Pt 7, p. 1488–1500, jul. 2014.

MCKEAN, Sandra C; DAVIES, John K; MOORE, Robert J. 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 (Reading, England)*, v. 153, n. Pt 7, p. 2203–11, jul. 2007a. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17600064>.

MCKEAN, Sandra C; DAVIES, John K; MOORE, Robert J. Probing the heat shock response of Corynebacterium pseudotuberculosis: the major virulence factor, phospholipase D, is downregulated at 43 degrees C. *Research in microbiology*, v. 158, n. 3, p. 279–86, abr. 2007b. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/17320354>.

MCNAMARA, P J; BRADLEY, G A; SONGER, J G. Targeted mutagenesis of the phospholipase D gene results in decreased virulence of Corynebacterium pseudotuberculosis. *Molecular microbiology*, v. 12, n. 6, p. 921–30, jun. 1994.

Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7934899>.

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. 113–8, 14 abr. 1995. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7737503>.

MEIBOM, Karin L. *et al.* The heat-shock protein ClpB of Francisella tularensis is involved in stress tolerance and is required for multiplication in target organs of infected mice. *Molecular Microbiology*, v. 67, n. 6, p. 1384–1401, mar. 2008.

MORITA, Y. *et al.* NorM, a putative multidrug efflux protein, of Vibrio parahaemolyticus and its homolog in Escherichia coli. *Antimicrobial Agents and Chemotherapy*, v. 42, n. 7, p. 1778–1782, jul. 1998.

MOTOHASHI, K. *et al.* Heat-inactivated proteins are rescued by the DnaK.J-GrpE set and ClpB chaperones. *Proceedings of the National Academy of Sciences of the United States of America*, v. 96, n. 13, p. 7184–7189, jun. 1999.

MOTTA, RODRIGO GARCIA; CREMASCO, Arita De Cássia Marella. Infecções por Corynebacterium pseudotuberculosis em animais de produção. *Veterinária e Zootecnia*, v. 17, n. 2, p. 200–213, 2010.

NAIR, Sudha; FINKEL, Steven E. Dps protects cells against multiple stresses during stationary phase. *Journal of Bacteriology*, v. 186, n. 13, p. 4192–4198, jul. 2004.

NEIDHARDT, Frederick C. (Org.). *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2 edition ed. Washington, D.C.: ASM Press, 1996.

NISHIMURA, Taku *et al.* Regulation of the nitrate reductase operon narKGHJI by the cAMP-dependent regulator GlxR in Corynebacterium glutamicum. *Microbiology (Reading, England)*, v. 157, n. Pt 1, p. 21–28, jan. 2011.

NOBLES, Christopher L; MARESSO, Anthony W. The theft of host heme by Gram-positive pathogenic bacteria. *Metallomics : integrated biometal science*, v. 3, n. 8, p. 788–96, ago. 2011. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21725569>.

OCHSNER, U A; JOHNSON, Z; VASIL, M L. Genetics and regulation of two distinct haem-uptake systems, phu and has, in Pseudomonas aeruginosa. *Microbiology (Reading, England)*, v. 146 (Pt 1, p. 185–98, jan. 2000. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10658665>.

OLSON, Merle E. *et al.* Biofilm bacteria: formation and comparative susceptibility to antibiotics. *Canadian Journal of Veterinary Research = Revue Canadienne De Recherche Vétérinaire*, v. 66, n. 2, p. 86–92, abr. 2002.

OSMUNDSON, Joseph; DEWELL, Scott; DARST, Seth A. RNA-Seq Reveals Differential Gene Expression in Staphylococcus aureus with Single-Nucleotide Resolution. *PLOS ONE*, v. 8, n. 10, p. e76572, jul. 2013.

PACHECO. *Caracterização do Exoproteoma Basal e Variante de Corynebacterium pseudotuberculosis e a Relação com a Virulência Reduzida de uma Linhagem Mutant*. 2010. 193 f. Universidade Federal de Minas Gerais, 2010.

PACHECO, Luis G. C. *et al.* A Role for Sigma Factor σE in Corynebacterium pseudotuberculosis Resistance to Nitric Oxide/Peroxide Stress. *Frontiers in*

Microbiology, v. 3, 2012.

PARROW, Nermi L.; FLEMING, Robert E.; MINNICK, Michael F. Sequestration and scavenging of iron in infection. *Infection and Immunity*, v. 81, n. 10, p. 3503–3514, 2013.

PATERSON, Gavin K.; MITCHELL, Tim J. The role of Streptococcus pneumoniae sortase A in colonisation and pathogenesis. *Microbes and Infection / Institut Pasteur*, v. 8, n. 1, p. 145–153, jan. 2006.

PEEL, M. M. *et al.* Human lymphadenitis due to Corynebacterium pseudotuberculosis: report of ten cases from Australia and review. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, v. 24, n. 2, p. 185–191, fev. 1997.

PÉPIN, M *et al.* Cellular composition of Corynebacterium pseudotuberculosis pyogranulomas in sheep. *Journal of leukocyte biology*, v. 56, n. 5, p. 666–70, nov. 1994. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7964175>.

PINTO, Anne *et al.* Differential transcriptional profile of Corynebacterium pseudotuberculosis in response to abiotic stresses. *BMC Genomics*, v. 15, n. 1, p. 14, 2014. Disponível em: <http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-14>.

PINTO, Anne Cybelle. Análise em larga escala da expressão diferencial de Corynebacterium pseudotuberculosis em resposta a estresses abióticos. p. 214, 2011.

PINTO, Anne Cybelle *et al.* The core stimulon of Corynebacterium pseudotuberculosis strain 1002 identified using ab initio methodologies. *Integrative Biology: Quantitative Biosciences from Nano to Macro*, v. 4, n. 7, p. 789–794, jul. 2012.

POTTER, Adam J. *et al.* Thioredoxin reductase is essential for protection of Neisseria gonorrhoeae against killing by nitric oxide and for bacterial growth during interaction with cervical epithelial cells. *The Journal of Infectious Diseases*, v. 199, n. 2, p. 227–235, jan. 2009.

RADOSTITS, O. M *et al. Clínica veterinária: um tratado de doenças dos bovinos, ovinos, suínos, caprinos e eqüinos*. Rio de Janeiro: Guanabara Koogan, 2002.

RAMOS, Juan L. *et al.* The TetR family of transcriptional repressors. *Microbiology and molecular biology reviews: MMBR*, v. 69, n. 2, p. 326–356, jun. 2005.

REEN, F. Jerry *et al.* A non-classical LysR-type transcriptional regulator PA2206 is required for an effective oxidative stress response in Pseudomonas aeruginosa. *PloS One*, v. 8, n. 1, p. e54479, 2013.

RIBEIRO, Dayana *et al.* An iron-acquisition-deficient mutant of Corynebacterium pseudotuberculosis efficiently protects mice against challenge. *Veterinary Research*, v. 45, n. 1, p. 28, 2014. Disponível em: <http://www.veterinaryresearch.org/content/45/1/28>.

RODRIGUE, Sébastien *et al.* The sigma factors of Mycobacterium tuberculosis. *FEMS microbiology reviews*, v. 30, n. 6, p. 926–941, nov. 2006.

RUIZ, Jerônimo C. *et al.* Evidence for reductive genome evolution and lateral acquisition of virulence functions in two Corynebacterium pseudotuberculosis strains.

PloS One, v. 6, n. 4, p. e18551, 2011.

RUTHERFORD, K. *et al.* Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)*, v. 16, n. 10, p. 944–945, out. 2000.

SANSOM, Fiona M. *et al.* Leishmania major methionine sulfoxide reductase A is required for resistance to oxidative stress and efficient replication in macrophages. *PloS One*, v. 8, n. 2, p. e56064, 2013.

SASINDRAN, Smitha J.; SAIKOLAPPAN, Sankaralingam; DHANDAYUTHAPANI, Subramanian. Methionine sulfoxide reductases and virulence of bacterial pathogens. *Future Microbiology*, v. 2, n. 6, p. 619–630, dez. 2007.

SCHUMANN, Wolfgang. *Functional Analysis of Bacterial Genes: A Practical Manual*. Sussex, United Kingdom: John Wiley e Sons, 2001.

SCHUMANN, Wolfgang. Thermosensors in eubacteria: role and evolution. *Journal of Biosciences*, v. 32, n. 3, p. 549–557, abr. 2007.

SELVY, Paige E *et al.* Phospholipase D: enzymology, functionality, and chemical modulation. *Chemical reviews*, v. 111, n. 10, p. 6064–119, 12 out. 2011. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21936578>.

SENDLER, Edward; JOHNSON, Graham D.; KRAWETZ, Stephen A. Local and global factors affecting RNA sequencing analysis. *Analytical Biochemistry*, v. 419, n. 2, p. 317–322, 2011.

SHELDON, Jessica R; LAAKSO, Holly A; HEINRICHS, David E. Iron acquisition strategies of bacterial pathogens. *Virulence mechanisms of bacterial pathogens*, n. May 2016, p. In press., 2016.

SHIBER, Ayala; RAVID, Tommer. Chaperoning proteins for destruction: diverse roles of Hsp70 chaperones and their co-chaperones in targeting misfolded proteins to the proteasome. *Biomolecules*, v. 4, n. 3, p. 704–724, 2014.

SHIN, Dong-Ho; CHOI, Young-Seok; CHO, You-Hee. Unusual properties of catalase A (KatA) of Pseudomonas aeruginosa PA14 are associated with its biofilm peroxide resistance. *Journal of Bacteriology*, v. 190, n. 8, p. 2663–2670, abr. 2008.

SHINNICK, T. M.; KING, C. H.; QUINN, F. D. Molecular biology, virulence, and pathogenicity of mycobacteria. *The American Journal of the Medical Sciences*, v. 309, n. 2, p. 92–98, fev. 1995.

SILVA, Wanderson M.; SEYFFERT, Núbia; CIPRANDI, Alessandra; *et al.* Differential Exoproteome Analysis of Two Corynebacterium pseudotuberculosis Biovar Ovis Strains Isolated from Goat (1002) and Sheep (C231). *Current Microbiology*, v. 67, n. 4, p. 460–465, out. 2013.

SILVA, Wanderson M.; SEYFFERT, Núbia; SANTOS, Agenor V.; *et al.* Identification of 11 new exoproteins in Corynebacterium pseudotuberculosis by comparative analysis of the exoproteome. *Microbial Pathogenesis*, v. 61–62, p. 37–42, ago. 2013.

SKAAR, Eric P. The battle for iron between bacterial pathogens and their vertebrate hosts. *PLoS pathogens*, v. 6, n. 8, p. e1000949, 2010. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20711357>.

SKOOG, WEST, HOLLER, Crouch. *FUNDAMENTOS DA QUÍMICA ANALÍTICA*. 8.

ed. [S.l: s.n.], 2005.

SMITH, Anthony W. Biofilms and antibiotic therapy: is there a role for combating bacterial resistance by the use of novel drug delivery systems? *Advanced Drug Delivery Reviews*, v. 57, n. 10, p. 1539–1550, jul. 2005.

SMITH, James L. The physiological role of ferritin-like compounds in bacteria. *Critical reviews in microbiology*, v. 30, n. 3, p. 173–85, 2004. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15490969>.

SOARES, Siomar C. *et al.* PIPS: pathogenicity island prediction software. *PloS One*, v. 7, n. 2, p. e30848, 2012.

SOARES, Siomar C. *et al.* The pan-genome of the animal pathogen Corynebacterium pseudotuberculosis reveals differences in genome plasticity between the biovar ovis and equi strains. *PloS One*, v. 8, n. 1, p. e53818, 2013.

SOCIETY, Royal Statistical. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing Author (s): Yoav Benjamini and Yosef Hochberg Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 57 , No . 1 Published by : . v. 57, n. 1, p. 289–300, 2016.

SQUIRES, C. L. *et al.* ClpB is the Escherichia coli heat shock protein F84.1. *Journal of Bacteriology*, v. 173, n. 14, p. 4254–4262, jul. 1991.

STRAUB, Shannon C.K. *et al.* Horizontal Transfer of DNA from the Mitochondrial to the Plastid Genome and Its Subsequent Evolution in Milkweeds (Apocynaceae). *Genome Biology and Evolution*, v. 5, n. 10, p. 1872–1885, out. 2013. Disponível em: <https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evt140>.

SUN, Ronggai *et al.* Mycobacterium tuberculosis ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set. *Molecular Microbiology*, v. 52, n. 1, p. 25–38, abr. 2004.

TAN, Mai Ping *et al.* Nitrate Respiration Protects Hypoxic Mycobacterium tuberculosis Against Acid- and Reactive Nitrogen Species Stresses. *PLoS ONE*, v. 5, n. 10, p. e13356, out. 2010.

TAO, Tao *et al.* Transcriptome Sequencing and Differential Gene Expression Analysis of Delayed Gland Morphogenesis in Gossypium australe during Seed Germination. *PLoS ONE*, v. 8, n. 9, p. e75323, 20 set. 2013. Disponível em: <http://dx.plos.org/10.1371/journal.pone.0075323>.

TASHJIAN, J J; CAMPBELL, S G. Interaction between caprine macrophages and corynebacterium pseudotuberculosis: an electron microscopic study. *American journal of veterinary research*, v. 44, n. 4, p. 690–3, abr. 1983. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/6869967>.

TONG, Yong; GUO, Maolin. Bacterial heme-transport proteins and their heme-coordination modes. *Archives of biochemistry and biophysics*, v. 481, n. 1, p. 1–15, 1 jan. 2009. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18977196>.

TRAPNELL, Cole *et al.* Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nature biotechnology*, v. 28, n. 5, p. 511–515, maio 2010.

TROST, Eva *et al.* Complete genome sequence and lifestyle of black-pigmented Corynebacterium aurimucosum ATCC 700975 (formerly C. nigricans CN-1) isolated from a vaginal swab of a woman with spontaneous abortion. *BMC genomics*, v. 11, p. 91, 2010. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20137072>.

TULLIUS, Michael V *et al.* Discovery and characterization of a unique mycobacterial heme acquisition system. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 12, p. 5051–5056, 2011. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21383189>.

VANNI, Stefano *et al.* Ion binding and internal hydration in the multidrug resistance secondary active transporter NorM investigated by molecular dynamics simulations. *Biochemistry*, v. 51, n. 6, p. 1281–1287, fev. 2012.

WEINBERG, E. D. Microbial pathogens with impaired ability to acquire host iron. 2000.

WENNERHOLD, Julia; BOTT, Michael. The DtxR Regulon of Corynebacterium glutamicum. *Journal of Bacteriology*, v. 188, n. 8, p. 2907–2918, 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1446976/>.

WILLIAMSON, L H. Caseous lymphadenitis in small ruminants. *The Veterinary clinics of North America. Food animal practice*, v. 17, n. 2, p. 359–71, vii, jul. 2001. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/11515406>.

YUAN, Yongxian; XU, Huaiqian; LEUNG, Ross Ka-Kit. An optimized protocol for generation and analysis of Ion Proton sequencing reads for RNA-Seq. *BMC Genomics*, v. 17, n. 1, p. 403, 26 dez. 2016. Disponível em: <http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2745-8>.

ZHANG, Wenli *et al.* Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *The International Journal of Biochemistry & Cell Biology*, v. 40, n. 11, p. 2560–2571, 2008.

ZHAO, Heng *et al.* Depletion of Undecaprenyl Pyrophosphate Phosphatases Disrupts Cell Envelope Biogenesis in Bacillus subtilis. *Journal of Bacteriology*, v. 198, n. 21, p. 2925–2935, nov. 2016.

ZOLKIEWSKI, M. ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from Escherichia coli. *The Journal of Biological Chemistry*, v. 274, n. 40, p. 28083–28086, out. 1999.

8. ANEXOS

8.1. ANEXO I: Material suplementar do artigo apresentado na seção 4.8.

Conteúdo:

Additional file 1: Table S1. Sequencing statistics

Additional file 1: Table S2. Total number of processed reads from High Iron (HI) and Low Iron (LI) experimental condition

Additional file 1: Table S3. Mapping results

Additional file 1: Table S4. Raw gene counts

Additional file 1: Figure S1. Sample to sample principal component analysis (PCA)

Additional file 2: Figure S2. *C. pseudotuberculosis* growth rate

Additional file 2: Figure S3. Effects of ethanol on bacterial growth

Additional file 3: Table S5. Differentially expressed genes identified in the T1 strain in response to iron limitation

Additional file 3: Table S6. Differentially expressed genes identified in the Cp13 mutant in response to iron limitation

Additional file 4: Table S7. Table of gene ontology terms assigned to DEGs identified in the T1 strain

Additional file 4: Table S8. Table of gene ontology terms assigned to DEGs identified in the Cp13 mutant

Additional file 5: Figure S4. Alignment of the amino acid sequence of the HtaA domain from protein products of the upregulated genes *htaC*, *htaA*, *htaF*, *htaG* and *Cp_3070* genes of *C. pseudotuberculosis*

Additional file 5: Figure S5. Structural protein domain characteristics of the HtaA, HtaC, HtaF, HtaG and Cp_3070 proteins

Additional file 5: Figure S6. Alignment of the amino acid sequences of the HmuU protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)

Additional file 5: Figure S7. Alignment of the amino acid sequences of the HmuT protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd)

Additional file 5: Figure S8. Alignment of the amino acid sequences of the DNA-binding regulator *hrrA* of the two-component regulatory system *hrrSA* between *C. pseudotuberculosis* (Cp), *C. glutamicum* (Cg) and *C. diphtheriae* (Cd)

Additional file 5: Table S9. Genomic Islands prediction

Additional file 1: Table S1. Sequencing statistics. Paired samples from High Iron (HI) and Low Iron (LI) experimental condition in the wild-type T1 and Cp13 strains were sequenced in two Ion proton runs (run 1 and run 2), yielding over 120 million single-end reads (124.251.275 total). Overall read length vary from 8 to 380 bp with an average CG content of 53-55%. The pair sample (S.3) from the wild-type T1 strain was eliminated from further analysis due to the low sequenced read count.

Strain	Sample ID	Condition	Average read length (bp)	Average CG content (%)	Number of reads	Batch Id	Run 1	Run 2
Cp13 mutant	S.1	HI	8-369	53	6,014,333	A	✓	
	S.1	LI	8-375	53	10,905,772	A	✓	
	S.2	HI_1	8-368	55	3,936,728	B	✓	
		HI_2	8-366	55	9,035,553	B		✓
	S.2	LI	8-368	54	4,834,840	C	✓	
	S.3	HI	8-374	55	8,974,704	C	✓	
	S.3	LI	8-378	55	4,236,781	C	✓	
	S.4	HI_1	8-368	54	1,789,770	C	✓	
		HI_2	8-369	54	3,327,609	C		✓
	S.4	LI_1	8-369	55	2,792,977	C	✓	
		LI_2	8-364	55	6,038,287	C		✓
Wild-type T1	S.1	HI_1	8-378	55	3,988,317	C	✓	
		HI_2	8-368	55	7,886,336	C		✓
	S.1	LI_1	8-355	55	3,537,940	C	✓	
		LI_2	8-366	55	7,801,599	C		✓
	S.2	HI	8-370	55	11,975,500	C	✓	
	S.2	LI	8-369	55	14,575,386	C	✓	
	S.3	HI	8-380	54	38,385	C	✓	
	S.3	LI	8-345	54	12,560,458	C		✓

Additional file 1: Table S2. Total number of processed reads from High iron (HI) and Low Iron (LI) experimental condition in the wild-type T1 and Cp13 mutant strains. Total number of reads, number of processed reads and percentage of excluded reads after applying a quality filtering score of 10 (Phred score 10) over the average base quality of the sliding window using Trimmomatic. Trimmed reads were quality checked using the FASTQC and presented an average read quality score of 24-25.

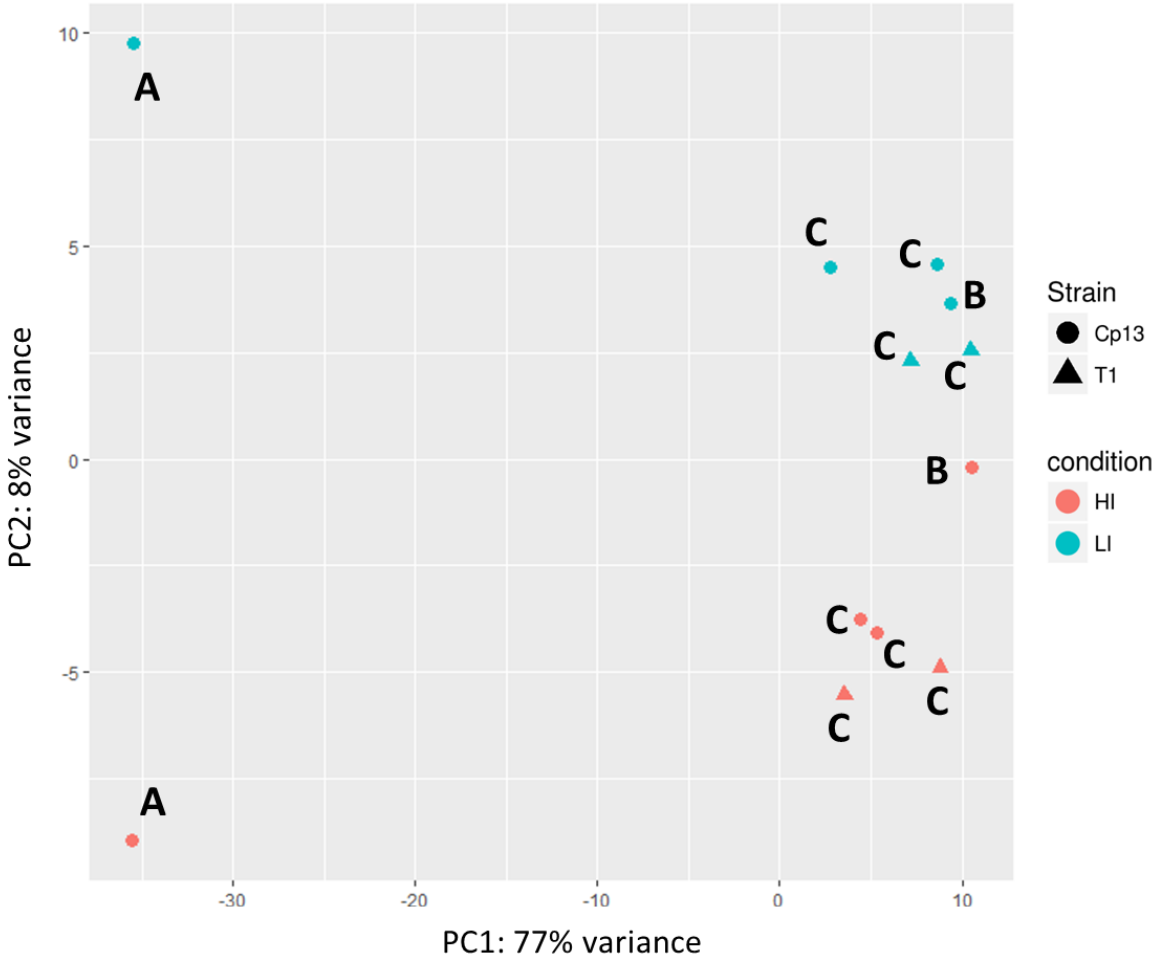
Strain	Sample Identification	Condition	Number of reads	Number of trimmed/processed reads	% of excluded reads	Phred score
Cp13 mutant	S.1	HI	6,014,33	6,012,724 (99.9%)	1,609 (0.03%)	24
	S.1	LI	10,905,772	10,902,511 (99.97%)	3,261 (0.03%)	24
	S.2	HI	3,936,728	3,931,106 (99.86%)	5,622 (0.14%)	25
			9,035,553	9,018,421 (99.81%)	17,132 (0,19%)	25
	S.2	LI	4,834,840	4,830,590 (99.91%)	4,250 (0.09%)	25
	S.3	HI	8,974,704	8.964,375 (99.88%)	10,329 (0.12%)	25
	S.3	LI	4,236,781	4,233,505 (99.92%)	3.276 (0.08%)	24
	S.4	HI	1,789,770	1,788,363 (99.92%)	1,407 (0.08%)	25
			3,327,609	3,324,827 (99.92%)	2,782 (0.08%)	24
	S.4	LI	2,792,977	2,791,946 (99.96%)	1,031 (0,04%)	25
			6,038,287	6,035.682 (99.96%)	2,605 (0.04%)	25
Wild-type T1	S.1	HI	3,988,317	3,985,364 (99.93%)	2,953 (0.07%)	25
			7,886,336	7,878,693 (99.90%)	7,643 (0.10%)	25
	S.1	LI	3,537,940	3,535,701 (99.94%)	2,239 (0.06%)	25
			7,801,599	7,795,681 (99.92%)	5,918 (0,08%)	25
	S.2	HI	11,975,500	11,967,330 (99.93%)	8,170 (0,07%)	25
	S.2	LI	14,575,386	14,561,588 (99.91%)	13,798 (0,09%)	24

Additional file 1: Table S3. Mapping results. Paired samples from High Iron (HI) and Low Iron (LI) experimental condition in the wild-type T1 and Cp13 mutant strains. The total number of reads and percentage of total single-end reads mapped to the *Corynebacterium pseudotuberculosis* strain 1002B genome (NZ_CP012837.1), using the Torrent Mapping Alignment Program (TMAP).

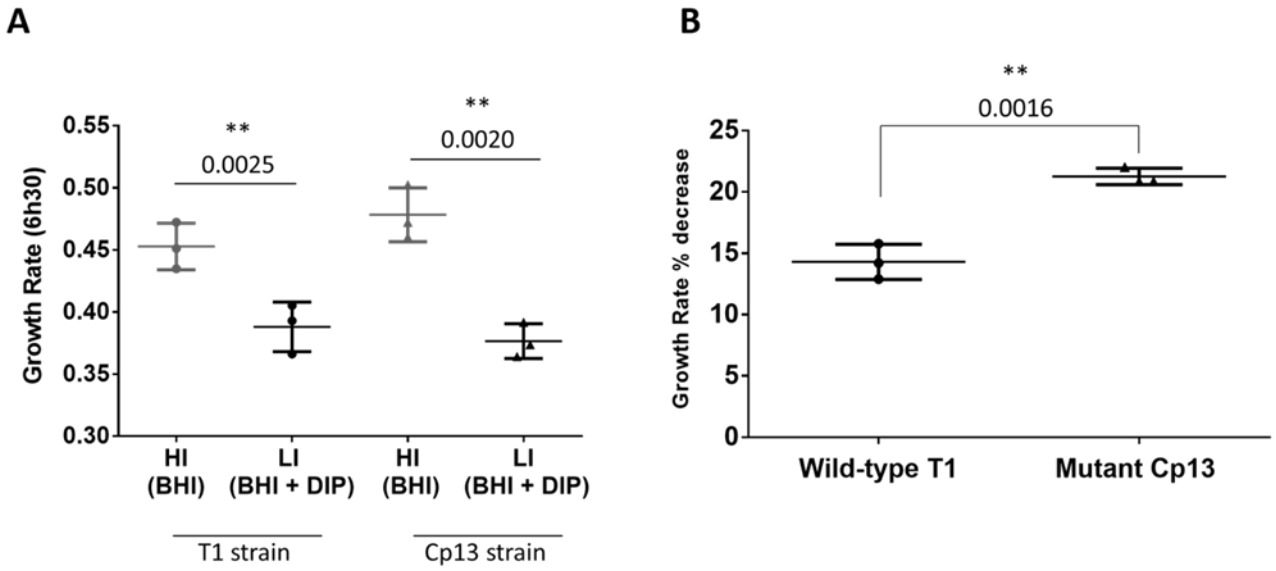
Strain	Sample Identification	Condition	Number of reads	Number of unmapped reads	Number of mapped reads	% of mapped reads
ciuA mutant Cp13	S.1	HI	6,012,724	77,775	5,934,949	98.7
	S.1	LI	10,902,511	117,242	10,785,269	98.9
	S.2	HI	12,949,527	3,232,235	9,717,292	75.0
	S.2	LI	4,830,590	1,325,346	3,505,244	72.6
	S.3	HI	8,964,375	1,776,069	7,188,306	80.2
	S.3	LI	4,233,505	672,232	3,561,273	84.1
	S.4	HI	5,113,190	935,598	4,177,592	81.7
	S.4	LI	8,827,628	2,348,088	6,479,540	73.4
Wild-type T1	S.1	HI	11,864,057	1,793,682	10,070,375	84.9
	S.1	LI	11,331,382	3,024,330	8,307,052	73.3
	S.2	HI	11,967,330	2,342,791	9,624,539	80.4
	S.2	LI	14,561,588	2,664,406	11,897,182	81.7

Additional file 1: Table S4. Number of unique and ambiguous (multiple) alignments. Paired samples from High Iron (HI) and Low Iron (LI) experimental condition in the wild-type T1 and Cp13 mutant strains with the total number of read counts aligned to unique features using HTSeq-count are presented in raw counted data and % of unique aligned reads. Reads that were assigned to more than one feature were counted as ambiguous and were not counted for any features (--nonunique none).

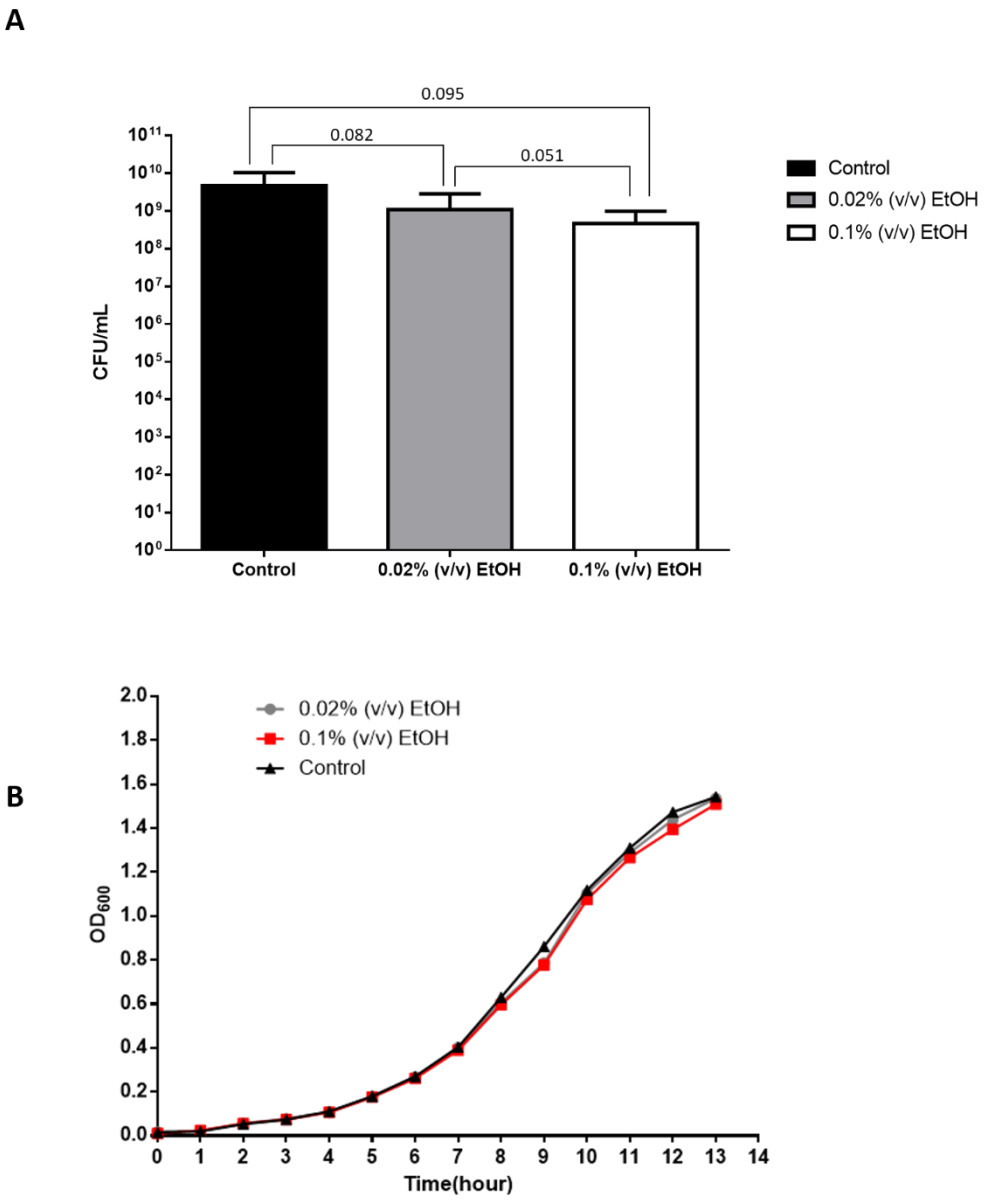
Strain	Sample Identification	Exp. Condition	Number of mapped reads	Ambiguous alignments	Not aligned	Unique alignments	% of unique alignments
ciuA mutant Cp13	S.1	HI	5,934,949	4,231,060	77,775	1,703,889	28.71
	S.1	LI	10,785,269	7,637,579	117,242	3,147,690	29.19
	S.2	HI	9,717,292	8,749,308	3,232,235	3,266,366	27.18
	S.2	LI	3,505,244	2,508,091	1,325,346	997,153	28.45
	S.3	HI	7,188,306	5,152,800	1,776,069	2,035,506	28.32
	S.3	LI	3,561,273	2,570,497	672,232	990,776	27.82
	S.4	HI	4,177,592	3,448,057	935,598	1,347,008	28.09
	S.4	LI	6,479,540	5,763,821	2,348,088	2,340,924	28.88
Wild-type T1	S.1	HI	10,070,375	8,060,346	1,793,682	3,232,915	28.63
	S.1	LI	8,307,052	7,426,007	3,024,330	2,994,533	28.74
	S.2	HI	9,624,539	6,900,236	2,342,791	2,724,303	28.31
	S.2	LI	11,897,182	8,622,027	2,664,406	3,275,155	27.53



Additional file 1: Figure S1. Sample to sample principal component analysis (PCA) and hierarchical clustering analyses. PCA plot for the biological replicates of the T1 wild-type and Cp13 mutant strains. PCA and heatmap were plotted with *rlog*-transformed data of normalized read counts. The first two principal components are shown on the X and Y axes, with the proportion of total variance attributed to that PC indicated. Each experimental sample is represented as a single point with color indicating experimental sample condition and shape indicating strain. Pink represents High Iron (HI) and blue Low Iron (LI) experimental samples. Strains were represented by geometrical figures: T1 wild-type strain (▲) and Cp13 mutant strain (●). Batches are labeled A-C. PC1 evidencing batch clustering and sample variation within batch A and batches B-C of the replicates. PC1 and PC2 represent 85% of the total variance of the samples. HI and LI grouping can be seen, representing 8% of the total variance (PC2). Batch effect was attribute to sample preparation and it was included into the DESeq2 design formula.



Additional file 2: Figure S2. *C. pseudotuberculosis* growth rate. Growth was calculated using the growth rate equation: $Growthrate(Gr) = log\ OD2 - log\ OD\ 1/(t_2 - t_1)$, where t_1 represents the optical density (OD_{600}) at the start of the incubation period and t_2 is the OD_{600} value at the final incubation time (6h30). Horizontal lines indicate the mean of 3 individual assays and vertical lines standard deviation (sd) from mean (**A**) Statistical analysis using a paired t-test confirmed that iron limitation induced a significant effect on the growth rate of the Cp13 (*p* value = 0.0020) and T1 strains (*p* value = 0.0025). No difference was observed when analyzing HI and LI samples between the strains (student t-test, *p* value > 0.05). (**B**) Limited iron availability induced an average decrease of 21.3% in the growth rate of the Cp13 mutant and 14.3% mean average decrease in the wild-type T1 strain.



Additional file 2: Figure S3. Effects of ethanol on bacterial growth. (**A**) The effects of ethanol on the growth of the *C. pseudotuberculosis* Cp13 strain was measured by determining the number of CFU/mL after 13 hours of incubation in BHI control medium and BHI media supplemented with 0,02% and 0,1% of ethanol (v/v). No significant difference was observed between the CFU counts of control and ethanol supplemented cultures ($p > 0.05$). (**B)** growth proliferation was measured hourly by optical density for 13 hours at 600nm in control, 0.02% and 0.01% of ethanol cultures.

Additional file 3:Table S5. Differentially expressed genes identified in the T1 strain in response to iron limitation

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013240877.1	*Cp1002B_95*	Hypothetical protein	1.70	-0.761910	2.14719E-05	DOWN
WP_013240878.1	*ppiA*	Peptidyl-prolyl cis-trans isomerase	2.07	-1.049914	1.85484E-06	DOWN
WP_013240933.1	*Cp1002B_350*	Transcription factor rok	1.54	-0.620500	1.32554E-07	DOWN
WP_013240983.1	*sdcS*	Sodium-dependent dicarboxylate transporter	2.35	-1.231800	2.83533E-18	DOWN
WP_013241065.1	*glxR*	Crp/Fnr family transcriptional regulator	1.59	0.668724	0.00295783	UP
WP_013241099.1	*lpd*	Dihydrolipoamide dehydrogenase	1.77	-0.821935	3.89617E-05	DOWN
WP_013241101.1	*sdhC*	Succinate dehydrogenase cytochrome b556 subunit	1.57	-0.649583	1.34394E-06	DOWN
WP_013241102.1	*sdhA*	Succinate dehydrogenase flavoprotein subunit	1.51	-0.590974	0.000102809	DOWN
WP_013241103.1	*sdhB*	Succinate dehydrogenase iron-sulfur subunit	1.86	-0.897827	6.39657E-15	DOWN
WP_013241104.1	*Cp1002B_1235*	Hypothetical protein	2.17	-1.120063	8.79442E-05	DOWN
WP_013241113.1	*Cp1002B_1280*	Hypothetical protein	1.36	-0.444455	0.002247494	REG
WP_013241147.1	*ccsA*	Cytochrome c biogenesis protein ccsa	1.53	0.617462	0.001022724	UP
WP_013241165.1	*rplA*	50S ribosomal protein L1	1.32	-0.403942	0.000666323	REG
WP_014300452.1	*rplJ*	50s ribosomal protein l10	1.74	-0.796356	0.000812793	DOWN
WP_013241171.1	*rplL*	50s ribosomal protein l7/l12	1.80	-0.845724	0.000186583	DOWN
WP_013241172.1	*Cp1002B_1590*	Hypothetical protein	1.37	-0.452040	0.000305149	REG
WP_013241174.1	*Cp1002B_1600*	Hypothetical protein	1.49	-0.577251	2.7149E-06	REG
WP_013241177.1	*Cp1002B_1615*	hypothetical protein	1.49	-0.570888	0.001660433	REG
WP_013241191.1	*rpsL*	30s ribosomal protein s12	1.86	-0.896705	3.80001E-06	DOWN
WP_013241193.1	*rpsL*	Elongation factor g (ef-g)	1.62	-0.695013	0.001030848	DOWN
WP_013241203.1	*rplC*	50s ribosomal protein l3	1.47	-0.553653	0.000169055	REG
WP_013241209.1	*rpsC*	30s ribosomal protein s3	1.50	-0.589148	2.82806E-06	DOWN
WP_013910826.1	*rpsM*	30S ribosomal protein S13	1.63	-0.705070	0.003009419	DOWN
WP_013241260.1	*rplM*	50S ribosomal protein L13	1.65	-0.719634	0.002412353	DOWN

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013241261.1	*rpsI*	30s ribosomal protein s9	1.97	-0.979783	2.95922E-07	DOWN
WP_013241263.1	*Cp1002B_2040*	Hypothetical protein	1.72	0.783698	0.001156371	UP
WP_013241266.1	*Cp1002B_2050*	hypothetical protein	1.67	0.741298	0.001724138	UP
WP_004566891.1	*groES*	Co-chaperone groes	1.70	-0.768056	7.83744E-06	DOWN
WP_013241276.1	*groEL*	molecular chaperone GroEL	1.86	-0.897579	0.001764785	DOWN
WP_013241310.1	*htaA*	Cell-surface hemin receptor	1.95	0.962387	2.56552E-14	UP
WP_013241311.1	*hmuT*	Hemin-binding periplasmic protein hmut	1.86	0.892215	1.20551E-13	UP
WP_013241312.1	*hmuU*	Hemin import atp-binding protein	1.50	0.584874	6.21045E-07	REG
WP_013241314.1	*htaC*	Membrane protein	6.93	2.793618	4.97068E-69	UP
WP_013241337.1	*accBC*	Acyl coenzyme a carboxylase	1.47	-0.551264	9.14909E-05	REG
WP_013241338.1	*cysA*	Sulfurtransferase	1.39	-0.472489	5.40782E-05	REG
WP_013241378.1	*Cp1002B_2580*	Hypothetical protein	1.50	-0.585806	0.000132477	DOWN
WP_014522151.1	*fecB*	Periplasmic binding protein	2.03	1.023079	1.84161E-10	UP
WP_013242872.1	*Cp1002B_2935*	Transglycosylase associated protein	2.47	-1.303432	3.28002E-17	DOWN
WP_013242865.1	*hrrS*	two-component system sensor histidine kinase	1.33	0.414158	0.003040232	REG
WP_013242864.1	*hrrA*	Two-component system transcriptional regulatory protein	1.68	0.747336	4.28669E-08	UP
WP_013242852.1	*sodA*	Manganese superoxide dismutase	1.94	-0.958009	5.64712E-16	DOWN
WP_014522292.1	*Cp1002B_3070*	Hypothetical protein	8.49	3.086410	3.5298E-37	UP
WP_013242841.1	*Cp1002B_3075*	Hypothetical protein	10.19	3.348728	2.60248E-50	UP
WP_014401453.1	*lytR*	Lytr family transcriptional regulator	1.47	0.553600	6.56646E-05	REG
WP_013242822.1	*glpQ*	glycerophosphoryl diester phosphodiesterase	1.38	-0.466029	0.000206749	REG
WP_013242563.1	*Cp1002B_4430*	Hypothetical protein	1.56	0.637332	4.6245E-07	UP
WP_013242548.1	*Cp1002B_4545*	Abc transporter atp-binding protein	1.36	-0.447958	0.001367634	REG
WP_013242547.1	*Cp1002B_4550*	Abc-type antimicrobial peptide transport system	1.55	-0.630213	0.000120287	DOWN
WP_014367533.1	*ftn*	Ferritin-like protein	2.36	-1.237090	7.85923E-15	DOWN
WP_013242523.1	*ctaD*	Cytochrome c oxidase subunit 1	1.93	-0.947923	1.54564E-07	DOWN
WP_013242515.1	*Cp1002B_4725*	Rhomboid family protein	1.66	-0.730178	4.04547E-05	DOWN

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013242461.1	*tig*	trigger factor	1.44	-0.525525	0.000524198	REG
WP_013242460.1	*clpP*	Atp-dependent clp protease proteolytic subunit	1.42	-0.507650	0.000343248	REG
WP_013242446.1	*rpmA*	50S ribosomal protein L27	1.51	-0.595384	0.001681922	DOWN
WP_014522238.1	*hmuO*	heme oxygenase	1.47	0.556247	0.00175222	REG
WP_013242332.1	*htaF*	Cell-surface hemin receptor	3.38	1.755780	4.15589E-27	UP
WP_013242331.1	*htaG*	Uncharacterized protein htac	2.78	1.473360	1.10176E-26	UP
WP_046341090.1	*aceF*	Dihydrolipoamide acyltransferase	2.25	-1.172387	8.6232E-09	DOWN
WP_013242285.1	*Cp1002B_5875*	hypothetical protein	1.42	-0.505116	0.002865682	REG
WP_013242283.1	*ctaC*	Cytochrome c oxidase subunit ii	2.48	-1.307544	4.1942E-15	DOWN
WP_013242282.1	*ctaF*	Cytochrome aa3 subunit 4	2.23	-1.157272	8.87325E-13	DOWN
WP_013242281.1	*ctaE*	Cytochrome c oxidase subunit iii	2.31	-1.205747	1.10736E-11	DOWN
WP_013242280.1	*qcrC*	Ubiquinol-cytochrome c reductase cytochrome c subunit	2.84	-1.508014	6.66346E-17	DOWN
WP_014300818.1	*qcrA*	Ubiquinol-cytochrome c reductase iron-sulfur subunit	2.35	-1.235498	1.50312E-06	DOWN
WP_013242278.1	*qcrB*	Ubiquinol-cytochrome c reductase cytochrome b subunit	2.42	-1.277399	3.57326E-15	DOWN
WP_013242277.1	*Cp1002B_5915*	Hypothetical protein, npl/p60 family rote	1.43	0.511751	0.000266284	REG
WP_013242261.1	*mraZ*	Transcriptional regulator mraz	1.95	-0.964893	3.28102E-07	DOWN
WP_013242245.1	*ag84*	Antigen 84	1.68	-0.744234	2.59512E-06	DOWN
WP_013242172.1	*xerC*	DNA processing protein DprA	1.30	0.383920	0.001755934	REG
WP_013242168.1	*tsf*	Elongation factor ts (ef-ts)	1.76	-0.819061	0.000609445	DOWN
WP_038616474.1	*mapB*	methionine aminopeptidase	1.40	0.485872	0.00149511	REG
WP_013242137.1	*ylxR*	DNA-binding protein	1.35	-0.432620	0.001920312	REG
WP_013242135.1	*rbfA*	Ribosome-binding factor a	1.62	-0.696887	5.74792E-05	DOWN
WP_013242094.1	*ptsH*	Phosphocarrier protein hpr	1.48	-0.568824	0.00131324	REG
WP_013242073.1	*sigB*	RNA polymerase sigma factor SigB	1.37	-0.457111	0.001568058	REG
WP_013242016.1	*citE*	citrate lyase subunit beta	1.43	0.513988	0.001650089	REG
WP_013241994.1	*nusB*	Transcription antitermination protein nusb	1.44	-0.526362	0.000136463	REG
WP_013241984.1	*Cp1002B_7360*	hypothetical protein	1.48	-0.569637	0.002657504	REG

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013241796.1	*doxX*	Doxx family protein	1.39	-0.472414	5.25994E-05	REG
WP_057061524.1	*Cp1002B_8515*	hypothetical protein	1.33	0.407664	0.002131544	REG
WP_014522179.1	*atpH*	F-type atpase subunit delta	1.58	-0.657731	0.000150468	DOWN
WP_013241701.1	*atpF*	F-type atpase subunit b	1.57	-0.654918	2.11641E-05	DOWN
WP_013241696.1	*ywlC*	Threonylcarbamoyl-amp synthase	1.60	0.678432	0.000725474	UP
WP_013241684.1	*lutB*	Lactate utilization protein b	2.11	-1.075741	3.23659E-15	DOWN
WP_013241683.1	*Cp1002B_8845*	Hypothetical protein	1.52	-0.605213	3.72574E-06	DOWN
WP_013241666.1	*galK*	galactokinase	1.29	0.366387	0.00275072	REG
WP_013241654.1	*odhA*	Alpha-ketoglutarate decarboxylase	1.45	-0.540864	0.002471747	REG
WP_013241593.1	*fumC*	Fumarate hydratase class ii (fumarase c) (ec 4.2.1.2) (aerobic fumarase) (iron-independent fumarase)	1.43	-0.511878	0.000272047	REG
WP_013241591.1	*Cp1002B_9300*	MFS transporter	1.32	0.401988	0.000330379	REG
WP_039697516.1	*ripA*	Arac family transcriptional regulator-hth-type transcriptional repressor of iron protein A	2.13	1.091910	1.19311E-12	UP
WP_013241535.1	*rpfB*	resuscitation-promoting factor	1.32	0.396790	0.002133833	REG
WP_013241520.1	*mscL*	Large conductance mechanosensitive channel protein mscl	1.49	-0.578272	0.000228706	REG
WP_013241517.1	*tcsS4*	two-component system sensor histidine kinase	1.45	-0.532569	0.000333247	REG
WP_013241513.1	*rpmB*	50s ribosomal protein l28	1.42	-0.509815	8.2349E-05	REG
WP_013241512.1	*rpsN*	30S ribosomal protein S14	1.44	-0.524718	0.001890195	REG
WP_013241487.1	*cobF*	precorrin 6A synthase	1.35	-0.436293	0.001649632	REG
WP_013241480.1	*arcB*	ornithine cyclodeaminase	1.47	0.557814	5.48628E-06	REG
WP_014300556.1	*Cp1002B_9880*	sodium:proton antiporter	1.39	0.475246	0.002782863	REG
WP_013241465.1	*fkbP*	Peptidyl-prolyl cis-trans isomerase	1.97	-0.980577	5.7635E-06	DOWN
WP_013241451.1	*rpfA*	Resuscitation-promoting factor	1.59	-0.665839	3.4754E-05	DOWN
WP_013242913.1	*Cp1002B_10210*	Hypothetical protein	5.43	-2.441716	1.17496E-19	DOWN
WP_013242924.1	*Cp1002B_10265*	Hypothetical protein	1.78	-0.827959	1.03231E-07	DOWN
WP_013240910.1	*Cp1002B_10495*	Hypothetical protein	4.97	2.313972	3.1878E-65	UP

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_003402602.1	*Cp1002B_10540*	AURKAIP1/COX24 domain-containing protein	1.84	-0.881351	0.002694085	DOWN
WP_014300599.1	*Cp1002B_10775*	Alkylmercury lyase	1.43	-0.515499	0.000454602	REG
WP_014300577.1	*Cp1002B_10785*	Hypothetical protein	3.59	1.843259	9.35463E-20	UP

Additional file 3:Table S6. Differentially expressed genes identified in the Cp13 mutant strain in response to iron limitation

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_014300357.1	*Cp1002B_70*	Hypothetical protein	2.12	1.08137012	1.69E-06	UP
WP_013240877.1	*Cp1002B_95*	Hypothetical protein	1.45	-0.53247980	0.019094153	REG
WP_013240983.1	*sdcS*	Sodium-dependent dicarboxylate transporter	1.47	-0.56035596	0.011372589	REG
WP_013240987.1	*Cp1002B_640*	Glyoxalase/Bleomycin resistance protein/Dihydroxybiphenyl dioxygenase	1.36	-0.44523800	0.014083724	REG
WP_013241021.1	*deoR*	Deor family transcriptional regulator	1.33	-0.40698491	0.025766871	REG
WP_013241084.1	*cspA*	Cold-shock protein	1.86	0.89665067	0.000775566	UP
WP_013241104.1	*Cp1002B_1235*	Hypothetical protein	1.99	-0.99080062	0.000258088	DOWN
WP_013241147.1	*ccsA*	Cytochrome c biogenesis protein ccsa	1.92	0.94304670	0.002894992	UP
WP_013241172.1	*Cp1002B_1590*	Membrane protein	1.33	-0.40746306	0.026664296	REG
WP_014300461.1	*Cp1002B_1680*	Hypothetical protein	1.86	-0.89842111	0.023503716	DOWN
WP_014401012.1	*Cp1002B_2240*	Membrane protein	1.35	0.42901158	0.022217473	REG
WP_013241310.1	*htaA*	Cell-surface hemin receptor	2.70	1.43267925	0.000128406	UP
WP_013241311.1	*hmuT*	ABC transporter substrate-binding protein	3.04	1.60250271	5.47E-05	UP
WP_013241312.1	*hmuU*	Hemin import atp-binding protein	2.09	1.06660938	0.005645516	UP
WP_013241314.1	*htaC*	Hypothetical protein	7.41	2.89015069	3.33E-11	UP
WP_014300514.1	*WhiB*	Whib family transcriptional regulator	1.59	0.66647696	0.020631043	UP
WP_013242872.1	*Cp1002B_2935*	Transglycosylase associated protein	1.75	-0.80851436	3.39E-05	DOWN
WP_013242864.1	*hrrA*	DNA-binding response regulator	1.99	0.98936897	4.02E-05	UP
WP_014522292.1	*Cp1002B_3070*	Hypothetical protein	22.90	4.51740393	3.62E-26	UP
WP_013242841.1	*Cp1002B_3075*	Hypothetical protein	28.20	4.81767729	2.43E-32	UP
	Cp1002B_3725	FUSC family protein	1.51	0.59106213	2.37E-05	UP
WP_014522271.1	*Cp1002B_3730*	Hypothetical protein	1.66	0.72793303	0.00335166	UP
WP_013242698.1	*ccsB*	Cytochrome C biogenesis protein	1.41	-0.49961433	0.044619839	REG
WP_013242669.1	*Cp1002B_3905*	Cation:proton antiporter	1.60	-0.67886689	0.02554289	DOWN
WP_013242602.1	*IlvB1*	Acetolactate synthase large subunit ilvb1	1.64	0.71442304	0.04042813	UP

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013242586.1	*Cp1002B_4315*	Hypothetical protein	1.94	0.95247557	0.001833731	UP
WP_014300921.1	*Cp1002B_4550*	Antimicrobial peptide ABC transporter	1.40	-0.48765514	0.003452582	REG
WP_013242430.1	*rpsT*	30S ribosomal protein S20	1.52	0.60224891	0.008322644	UP
WP_013242381.1	*Cp1002B_5410*	Membrane protein	1.38	0.46069178	0.037698195	REG
WP_013242364.1	*Cp1002B_5505*	Hypothetical protein	1.66	0.73058897	8.84E-06	UP
WP_013242345.1	*Cp1002B_5600*	Hypothetical protein	1.49	-0.57305480	0.028963703	REG
WP_013242332.1	*htaF*	Cell-surface hemin receptor	7.29	2.86676168	2.69E-14	UP
WP_013242331.1	*htaG*	Uncharacterized protein htac	5.86	2.55071095	0.002597513	UP
WP_013242283.1	*ctaC*	Cytochrome c oxidase subunit II	2.10	-1.06697640	1.06E-06	DOWN
WP_013242282.1	*ctaF*	Cytochrome c oxidase polypeptide 4	2.13	-1.08882811	1.08E-07	DOWN
WP_013242281.1	*ctaE*	Heme-copper oxidase subunit III;Cytochrome c oxidase subunit III	2.61	-1.38133384	7.24E-11	DOWN
WP_013242280.1	*qcrC*	Ubiquinol-cytochrome C reductase cytochrome C subunit	2.57	-1.36030789	1.02E-09	DOWN
WP_014300818.1	*qcrA*	Ubiquinol-cytochrome c reductase iron-sulfur subunit	2.18	-1.12369409	0.000567675	DOWN
WP_013242278.1	*qcrB*	Ubiquinol-cytochrome C reductase cytochrome B subunit	2.44	-1.28647712	6.48E-07	DOWN
WP_013242277.1	*Cp1002B_5915*	Hypothetical protein, npl/p60 family rote	2.10	1.06894276	1.59E-11	UP
WP_013242223.1	*Cp1002B_6180*	Hypothetical protein	1.51	0.59905564	2.43E-06	UP
WP_013242184.1	*rpsP*	30S ribosomal protein S16	1.68	0.75001067	0.006012873	UP
WP_013242169.1	*rpsB*	30S ribosomal protein S2	1.86	0.89254186	0.011455716	UP
WP_038616474.1	*mapB*	30S ribosomal protein S2	1.26	0.33579134	0.043912281	REG
WP_013242135.1	*rbfA*	Ribosome-binding factor A	1.54	-0.61916475	0.002597513	DOWN
WP_013242127.1	*rpsO*	30S ribosomal protein S15	1.66	0.72694547	0.000258088	UP
WP_013242121.1	*Cp1002B_6680*	Hypothetical protein	1.41	-0.49825912	0.018377372	REG
WP_013242080.1	*ahpD*	Alkyl hydroperoxide reductase	1.45	0.53549987	0.002212884	REG
WP_013241996.1	*Cp1002B_7300*	Uncharacterized peptidase yqht	1.36	0.44186931	0.045562918	REG
WP_013241994.1	*nusB*	Transcription antitermination protein nusb	1.31	0.39293454	0.044444147	REG

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013241968.1	*rapZ*	Rnase adaptor protein rapz	1.32	0.40055688	0.026878656	REG
WP_013241958.1	*zwf*	Glucose-6-phosphate dehydrogenase	1.41	0.49092795	0.006012873	REG
WP_013241957.1	*tal*	Transaldolase	1.27	0.34747651	0.023121339	REG
WP_013241950.1	*sufR*	Arsr family transcriptional regulator;Iron-sulfur cluster biosynthesis transcriptional regulator sufr	1.65	0.72331015	0.000883021	UP
WP_013241891.1	*pafB*	Protein pafb	1.54	0.62541924	0.001833731	UP
WP_013241870.1	*ndh*	NAD(P)/FAD-dependent oxidoreductase	1.58	0.66087849	0.002206551	UP
WP_014522824.1	*gnd*	Phosphogluconate dehydrogenase	1.43	0.51097812	0.00479576	REG
WP_013241858.1	*Cp1002B_7990*	Hypothetical protein	1.35	0.43456513	0.049631926	REG
WP_057061524.1	*Cp1002B_8515*	Hypothetical protein	1.55	0.63109337	0.014090059	UP
WP_038616439.1	*Cp1002B_8520*	Hypothetical protein	1.56	0.63873835	0.003003536	UP
WP_014300634.1	*czcD*	Cadmium, cobalt and zinc/H(+)-K(+) antiporter	1.98	0.98243851	2.57E-05	UP
WP_013241750.1	*ilvC*	Ketol-acid reductoisomerase (NADP(+))	1.42	0.50216481	0.026701234	REG
WP_013241696.1	*ywlC*	Trna threonylcarbamoyladenosine biosynthesis protein ywlc	1.52	0.60023873	3.39E-05	UP
WP_013241682.1	*Cp1002B_8855*	Hypothetical protein	1.95	0.96684071	9.02E-05	UP
WP_014366861.1	*GlpR*	Glycerol-3-phosphate regulon repressor	1.64	0.71631969	0.007008324	UP
WP_013241665.1	*galT*	Galactose-1-phosphate uridylyltransferase	1.45	0.53877654	0.003194642	REG
WP_013241622.1	*Cp1002B_9145*	Hypothetical protein	1.43	0.51586838	0.002894992	REG
WP_013241575.1	*Cp1002B_9380*	Hypothetical protein	1.74	0.80159961	1.05E-08	UP
WP_039697516.1	*ripA*	Arac family transcriptional regulator-HTH-type transcriptional repressor of iron protein A	1.74	1.28829471	6.08E-13	UP
WP_013241560.1	*Cp1002B_9460*	Ribose-phosphate pyrophosphokinase	1.33	-0.40905274	0.014460429	REG
WP_013241539.1	*Cp1002B_9550*	Antibiotic biosynthesis monooxygenase	1.51	0.59258290	0.004739101	UP
WP_013241535.1	*rpfB*	Resuscitation-promoting factor rpfb	1.70	0.76406592	1.58E-07	UP
WP_013241514.1	*rpmE2*	50S ribosomal protein L31	1.38	0.46891228	0.006379392	REG
WP_013241482.1	*Cp1002B_9845*	ABC transporter substrate-binding protein	1.47	0.56022364	0.013264456	REG
WP_013241465.1	*fkbP*	Peptidyl-prolyl cis-trans isomerase, FKBP-type	1.61	-0.68715082	0.01323332	DOWN

Refseq_protein	Gene/Locus ID	Product	LI:Ct Fold Change	LI:Ct Log2FoldChange	FDR	Gene reg.
WP_013242913.1	*Cp1002B_10210*	Substrate-binding protein	2.44	-1.28477155	1.28E-07	DOWN
WP_013240910.1	*Cp1002B_10495*	Hypothetical protein	6.30	2.65529734	2.66E-14	UP
WP_014300401.1	*Cp1002B_10510*	Hypothetical protein	1.80	0.85097217	0.013271231	UP
WP_076761485.1	*Cp1002B_10545*	Hypothetical protein	1.81	0.85570870	0.04264298	UP
WP_014300577.1	*Cp1002B_10785*	Hypothetical protein	5.93	2.56871970	3.23E-32	UP

Additional file 4: Table S6. Table of gene ontology terms assigned to DEGs identified in the T1 strain. Genes have been divided in up and downregulated and classified into biological process, cellular function and molecular function categories. The *hits* highlight the number of genes assigned to a specific GO term. The GO annotations for DEGs were submitted to GOfeat for ontology classification.

Gene ontology classification of differentially expressed genes in the T1 wild-type strain			
Upregulated genes identified in the T1 strain			
Biological process			
GO	GO term	Hits	gene ID
GO:0006520	cellular amino acid metabolic process	1	*Cp1002B_02285*
GO:0017004	cytochrome complex assembly	1	*ccsA*
GO:0000160	phosphorelay signal transduction system	1	*hrrA*
GO:0006355	regulation of transcription, DNA-templated	1	*hrrA*
GO:0006351	transcription, DNA-templated	2	*hrrA; ripA*
Cellular component			
GO:0016021	integral component of membrane	9	*ccsA; htaA; hmuU; htaC; htaG; htaF;Cp1002B_3075; lytR; Cp1002B_3070;*
GO:0005622	intracellular	1	*hrrA*
GO:0005886	plasma membrane	1	*hmuU*
Molecular function			
GO:0005524	ATP binding	1	*hmuU*
GO:0003677	DNA binding	1	*hrrA*
GO:0003700	DNA binding transcription factor activity	1	*ripA*
GO:0003725	double-stranded RNA binding	1	*ywlC*
GO:0020037	heme binding	1	*ccsA*
GO:0030170	pyridoxal phosphate binding	1	*Cp1002B_RS02285*
GO:0043565	sequence-specific DNA binding	1	*ripA*
GO:0016765	transferase activity, transferring alkyl or aryl (other than methyl) groups	1	*Cp1002B_RS02285*
GO:0005215	transporter activity	1	*hmuU*
Downregulated genes identified in the T1 strain			
Biological process			
GO:0009060	aerobic respiration	1	*ctaD*
GO:0015986	ATP synthesis coupled proton transport	2	*atpF; atpH*
GO:0007049	cell cycle	1	*ag84*
GO:0051301	cell division	1	*ag84*
GO:0045454	cell redox homeostasis	2	*lpd; trxA2*
GO:0006879	cellular iron ion homeostasis	1	*ftn*
GO:0006106	fumarate metabolic process	1	*fumC*
GO:0006662	glycerol ether metabolic process	1	*trxA2*
GO:0006826	iron ion transport	1	*ftn*
GO:0019516	lactate oxidation	1	*lutB*
GO:0030490	maturation of SSU-rRNA	1	*rbfA*

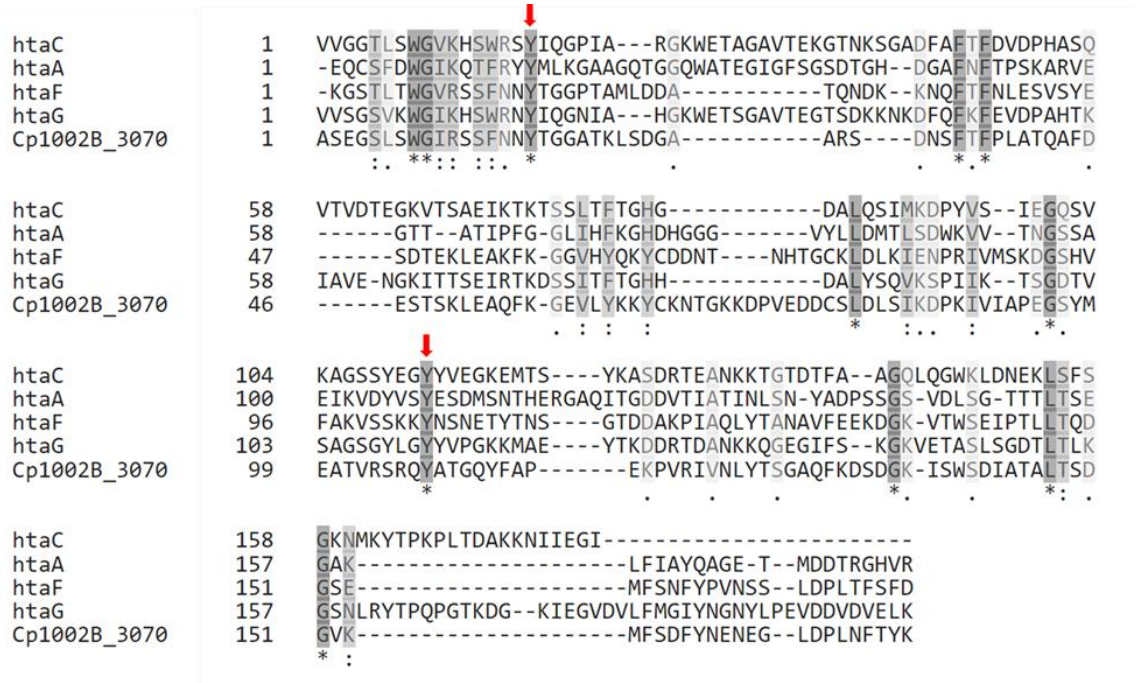
Gene ontology classification of differentially expressed genes in the T1 wild-type strain			
GO:0008152	metabolic process	1	*aceF*
GO:0006119	oxidative phosphorylation	1	*ctaD*
GO:0044238	primary metabolic process	1	*Cp1002B_02580*
GO:0006457	protein folding	1	*ppiA*
GO:0006417	regulation of translation	1	*Cp1002B_02580*
GO:0022904	respiratory electron transport chain	1	*qcrB*
GO:0042254	ribosome biogenesis	1	*rplJ*
GO:0006814	sodium ion transport	1	*sdcS*
GO:0006351	transcription, DNA-templated	2	*mraZ glxr*
GO:0006412	translation	5	*rplL; rpsL; rpsC; rpsl; rplJ*
GO:0055085	transmembrane transport	1	*sdcS*
GO:0006099	tricarboxylic acid cycle	3	*sdhB; fumC; Cp1002B_08990*
	Cellular component		
GO:0005623	cell	2	*lpd; trxA2*
GO:0005737	cytoplasm	8	*fusA; ptsH; rbfA; tsf; ag84; mraZ; clpP; ftn*
GO:0016021	integral component of membrane	17	*Cp1002B_00095; sdcS; sdhC; Cp1002B_01235; Cp1002B_01600; atpF; qcrB; qcrC; ctaE;ctaF; ctaC; Cp1002B_04725; ctaD; Cp1002B_04550; Cp1002B_02935; Cp1002B_10210; qcrA*
GO:0009295	nucleoid	1	*mrzA*
GO:0005886	plasma membrane	6	*atpF; ctaE; ctaF ctaD; Cp1002B_04550; atpH*
GO:0045261	proton-transporting ATP synthase complex, catalytic core F(1)	1	*atpH*
GO:0045263	proton-transporting ATP synthase complex, coupling factor F(o)	1	*atpF*
GO:0070469	respiratory chain	1	*ctaD*
GO:0005840	ribosome	5	*rplL;rpsl;Cp1002B_02580; rpsR; rplJ*
GO:0015935	small ribosomal subunit	2	*rpsL; rpsC*
GO:0045239	tricarboxylic acid cycle enzyme complex	1	*fumC*
	Molecular function		
GO:0051537	2 iron, 2 sulfur cluster binding	2	*sdhB; qcrA*
GO:0005524	ATP binding	2	*accBc; Cp1002B_04545*
GO:0016887	ATPase activity	1	*Cp1002B_04545*
GO:0004075	biotin carboxylase activity	1	*accBc*
GO:0005507	copper ion binding	2	*ctaC; ctaD*
GO:0004129	cytochrome-c oxidase activity	4	*ctaE; ctaF; ctaC; ctaD*
GO:0004148	dihydrolipoyl dehydrogenase activity	1	*lpd*
GO:0003677	DNA binding	2	*mraZ; vapI*

Gene ontology classification of differentially expressed genes in the T1 wild-type strain			
GO:0003700	DNA binding transcription factor activity	1	*mraZ*
GO:0009055	electron transfer activity	4	*lpd; sdhB; qcrB; qcrC*
GO:0008199	ferric iron binding	1	*ftn*
GO:0050660	flavin adenine dinucleotide binding	1	*lpd*
GO:0004333	fumarate hydratase activity	1	*fumC*
GO:0005525	GTP binding	1	*fusA*
GO:0003924	GTPase activity	1	*fusA*
GO:0020037	heme binding	2	*qcrC; ctaD*
GO:0005506	iron ion binding	2	*qcrC; ctaD*
GO:0070180	large ribosomal subunit rRNA binding	1	*rplJ*
GO:0046872	metal ion binding	2	*accBc; sodA*
GO:0003729	mRNA binding	1	*rpsC*
GO:0016491	oxidoreductase activity	1	*sdhA*
GO:0016679	oxidoreductase activity, acting on diphenols and related substances as donors	1	*qcrA*
GO:0016627	oxidoreductase activity, acting on the CH-CH group of donors	1	*sdhC*
GO:0004591	oxoglutarate dehydrogenase (succinyl-transferring) activity	1	*Cp1002B_08990*
GO:0003755	peptidyl-prolyl cis-trans isomerase activity	2	*ppiA; fkbP*
GO:0015035	protein disulfide oxidoreductase activity	1	*qcrC*
GO:0046933	proton-transporting ATP synthase activity, rotational mechanism	2	*atpF; atpH*
GO:0019843	rRNA binding	2	*rpsL; rpsC*
GO:0004252	serine-type endopeptidase activity	2	*clpP; Cp1002B_04725*
GO:0003735	structural constituent of ribosome	5	*rplL; rpsL; rpsC; rpsI; rplJ*
GO:0004784	superoxide dismutase activity	1	*sodA*
GO:0030976	thiamine pyrophosphate binding	1	*Cp1002B_08990*
GO:0004792	thiosulfate sulfurtransferase activity	1	*cysA*
GO:0016746	transferase activity, transferring acyl groups	2	*Cp1002B_08990; aceF*
GO:0003746	translation elongation factor activity	2	*fusA; tsf*
GO:0005215	transporter activity	1	*sdcS*
GO:0000049	tRNA binding	1	*rpsL*

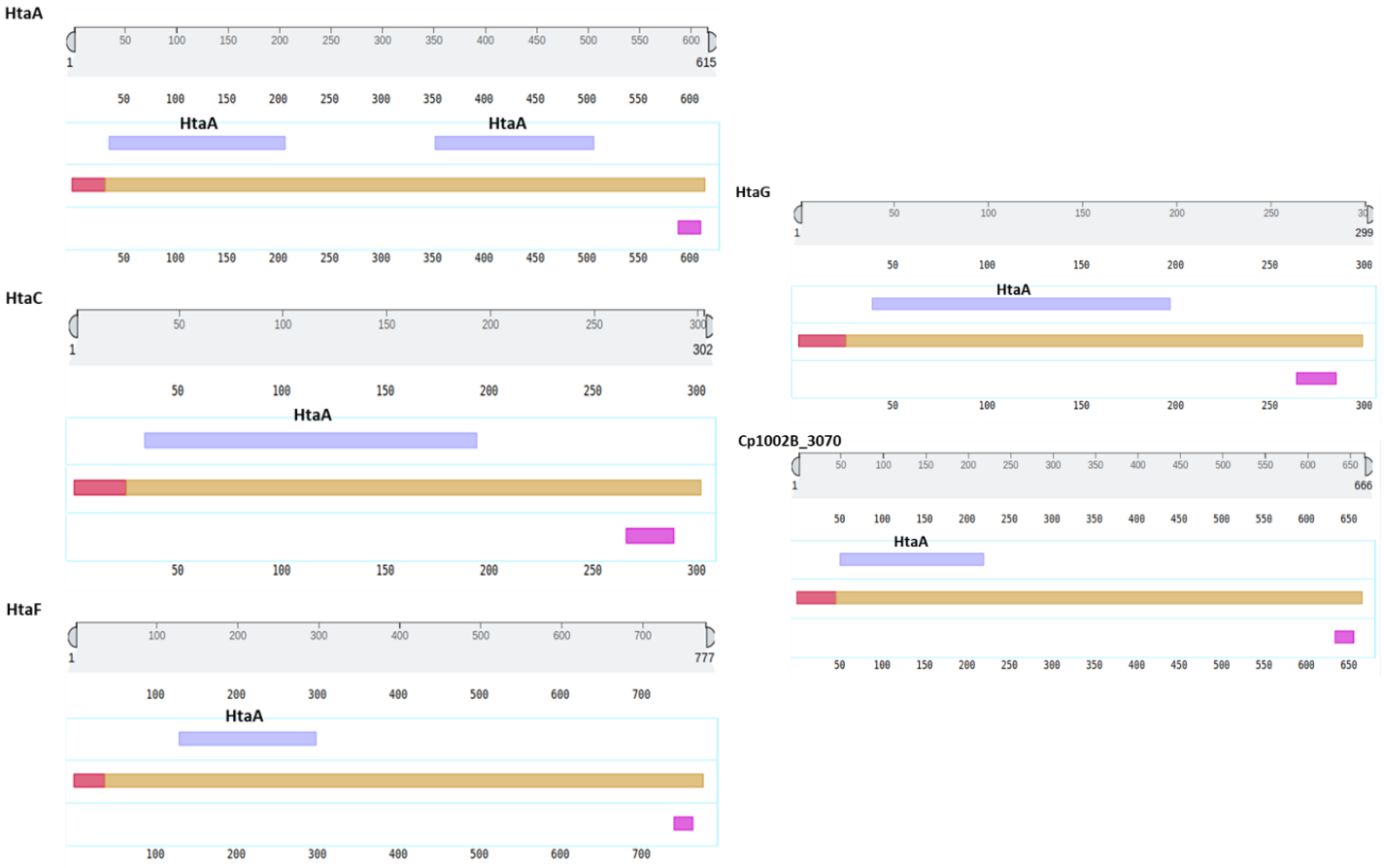
Additional file 4: Table S7. Table of gene ontology terms assigned to DEGs identified in the Cp13 mutant. Genes have been divided in up and downregulated and classified into the biological process, cellular function and molecular function categories. The *hits* highlight the number of genes assigned to a specific GO term. The GO annotations for DEGs were submitted to GOfeat for ontology classification.

Gene ontology classification of differentially expressed genes in the Cp13 mutant strain			
Up-regulated genes in Cp13			
Biological process			
GO	GO term	Hits	Gene ID
GO:0017004	cytochrome complex assembly	1	*ccsA*
GO:0009143	nucleoside triphosphate catabolic process	1	*mazG*
GO:0000160	phosphorelay signal transduction system	1	*hrrA*
GO:0006355	regulation of transcription, DNA-templated	3	*cspA; hrrA; whiB*
GO:0006351	transcription, DNA-templated	4	*hrrA; whiB; Cp1002B_08860; ripA*
GO:0006412	translation	4	*rpsO; rpsB; rpsP; rspT*
GO:0055085	transmembrane transport	1	*oppA2*
Cellular component			
GO:0043190	ATP-binding cassette (ABC) transporter complex	1	*oppA2*
GO:0005737	cytoplasm	2	*cspA; whiB*
GO:0016021	integral component of membrane	10	*ccsA; htaA; htaC; Cp1002B_09380; ndh; htaG; htaF; Cp1002B_03075; czcD; Cp1002B_03070*
GO:0005622	intracellular	2	*hrrA; Cp1002B_08860*
GO:0005840	ribosome	3	*rpsO; rpsP; rspT*
GO:0015935	small ribosomal subunit	1	*rpsB*
Molecular function			
GO:0051539	4 iron, 4 sulfur cluster binding	1	*whiB*
GO:0008324	cation transmembrane transporter activity	1	*czcD*
GO:0035731	dinitrosyl-iron complex binding	1	*whiB*
GO:0003677	DNA binding	4	*cspA; hrrA; whiB; Cp1002B_08860*
GO:0003700	DNA binding transcription factor activity	3	*sufR; Cp1002B_08860; ripA*
GO:0020037	heme binding	1	*ccsA*
GO:0046872	metal ion binding	2	*mazG; whiB*
GO:0004497	monooxygenase activity	1	*Cp1002B_09550*
GO:0047429	nucleoside-triphosphate diphosphatase activity	1	*mazG*
GO:0016491	oxidoreductase activity	1	*ndh*
GO:0019843	rRNA binding	2	*rpsO; rspT*
GO:0043565	sequence-specific DNA binding	1	*ripA*
GO:0046872	structural constituent of ribosome	4	*rpsO; rpsB; rpsP; rspT*
Downregulated genes identified in the Cp13 strain			

Gene ontology classification of differentially expressed genes in the Cp13 mutant strain			
Biological process			
GO:0019646	aerobic electron transport chain	1	*ctaE*
GO:0030490	maturation of SSU-rRNA	1	*rbfA*
GO:0022904	respiratory electron transport chain	1	*qcrB*
GO:0006814	sodium ion transport	1	*sdcS*
GO:0055085	transmembrane transport	1	*sdcS*
GO:0044206	UMP salvage	1	*upp*
GO:0006223	uracil salvage	1	*upp*
Cellular component			
GO:0005737	cytoplasm	1	*rbfA*
GO:0016021	integral component of membrane	10	*sdcS; Cp1002B_01235; qcrB; qcrC; ctaE; ctaF; ctaC; Cp1002B_02935; Cp1002B_10210; qcrA*
GO:0005886	plasma membrane	2	*ctaE; ctaF*
Molecular function			
GO:0051537	2 iron, 2 sulfur cluster binding	1	*qcrA*
GO:0005507	copper ion binding	1	*ctaC*
GO:0004129	cytochrome-c oxidase activity	3	*ctaE; ctaF; ctaC*
GO:0009055	electron transfer activity	2	*qcrB; qcrC*
GO:0005525	GTP binding	1	*upp*
GO:0020037	heme binding	1	*qcrC*
GO:0005506	iron ion binding	1	*qcrC*
GO:0000287	magnesium ion binding	1	*upp*
GO:0016679	oxidoreductase activity, acting on diphenols and related substances as donors	1	*qcrA*
GO:0003755	peptidyl-prolyl cis-trans isomerase activity	1	*fkbP*
GO:0005215	transporter activity	1	*sdcS*
GO:0004845	uracil phosphoribosyltransferase activity	1	*upp*

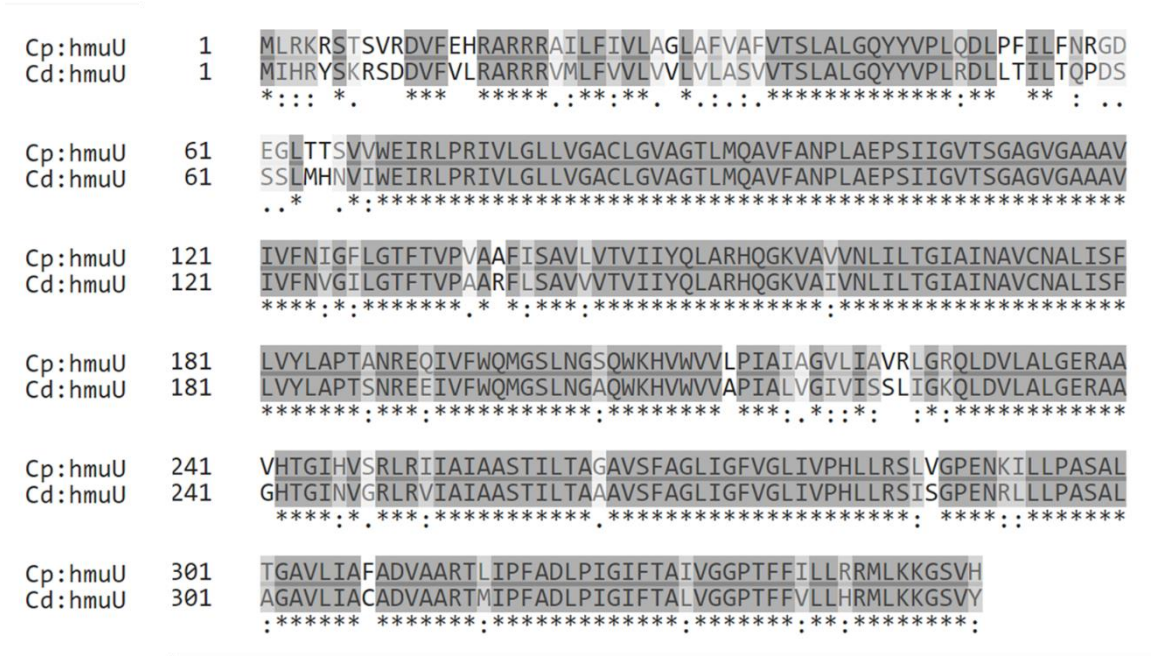


Additional file 5: Figure S4. Alignment of the amino acid sequence of the HtaA domain from protein products of the upregulated genes *htaC*, *htaA*, *htaF*, *htaG* and *Cp1002B_3070* genes of *C. pseudotuberculosis*. Conserved tyrosine residues are indicated above the sequence alignment with arrows; asterisks indicate sequence identity and similarity is indicated by colons and periods.

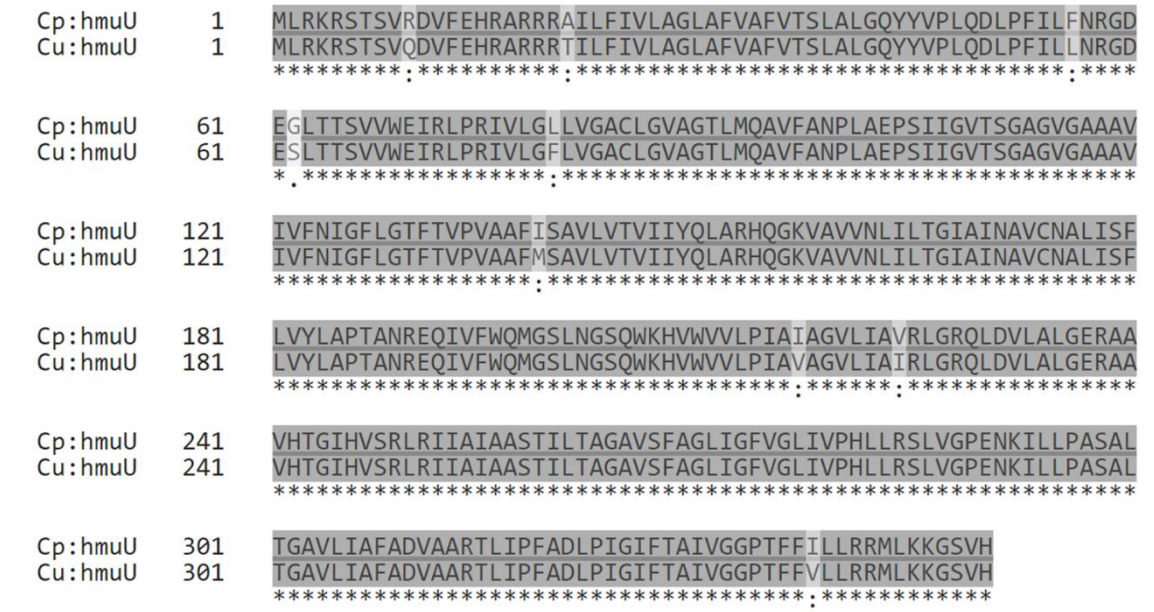


Additional file 5: Figure S5. Structural protein domain characteristics of the HtaA, HtaC, HtaF, HtaG and Cp_3070 proteins. The N-terminal signal peptide, binding HtaA and C-terminal transmembrane domains are schematized by color: signal peptide (red), htaA (purple) and transmembrane (pink). The predicted proteins contain at least one HtaA domain (two for the HtaA protein) consisting of approximately 150 amino acids with a conserved hemin and hemoglobin binding tyrosine residue, a C-terminal transmembrane domain of approximately 22 amino acids responsible for anchoring these proteins to the cytoplasmic membrane with predicted signal peptides.

A

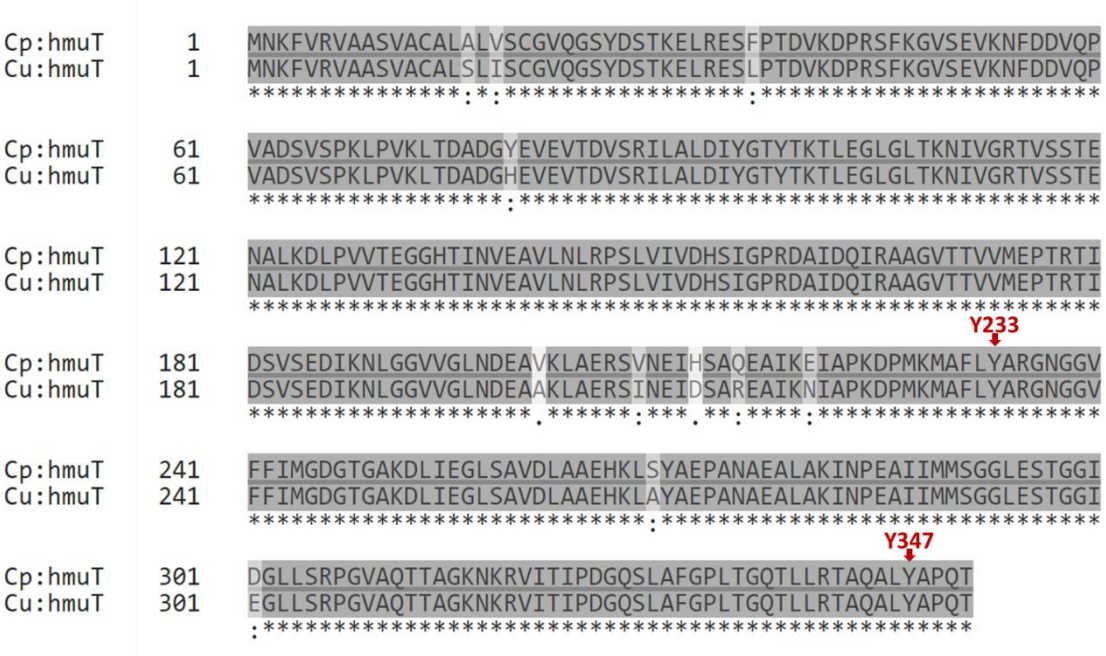


B

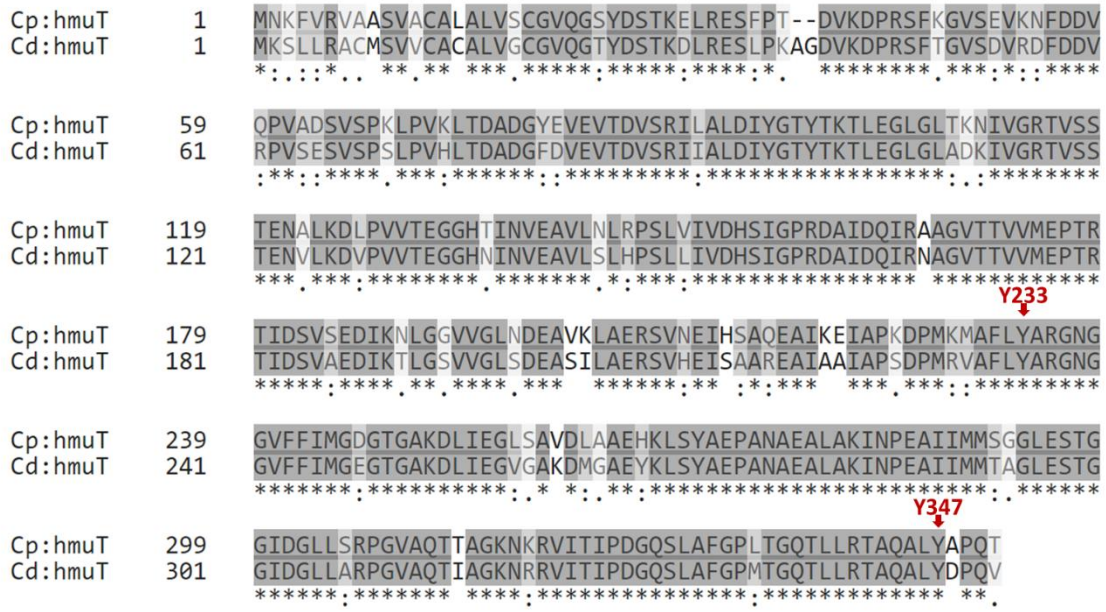


Additional file 5: Figure S6. Alignment of the amino acid sequences of the HmuU protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd). Asterisks indicate sequence identity and similarity is indicated by colons and period. (**A**) Cp and Cu alignment with 97.42% amino acid sequence identity; (**B**) Cp and Cd alignment with 80% amino acid sequence identity.

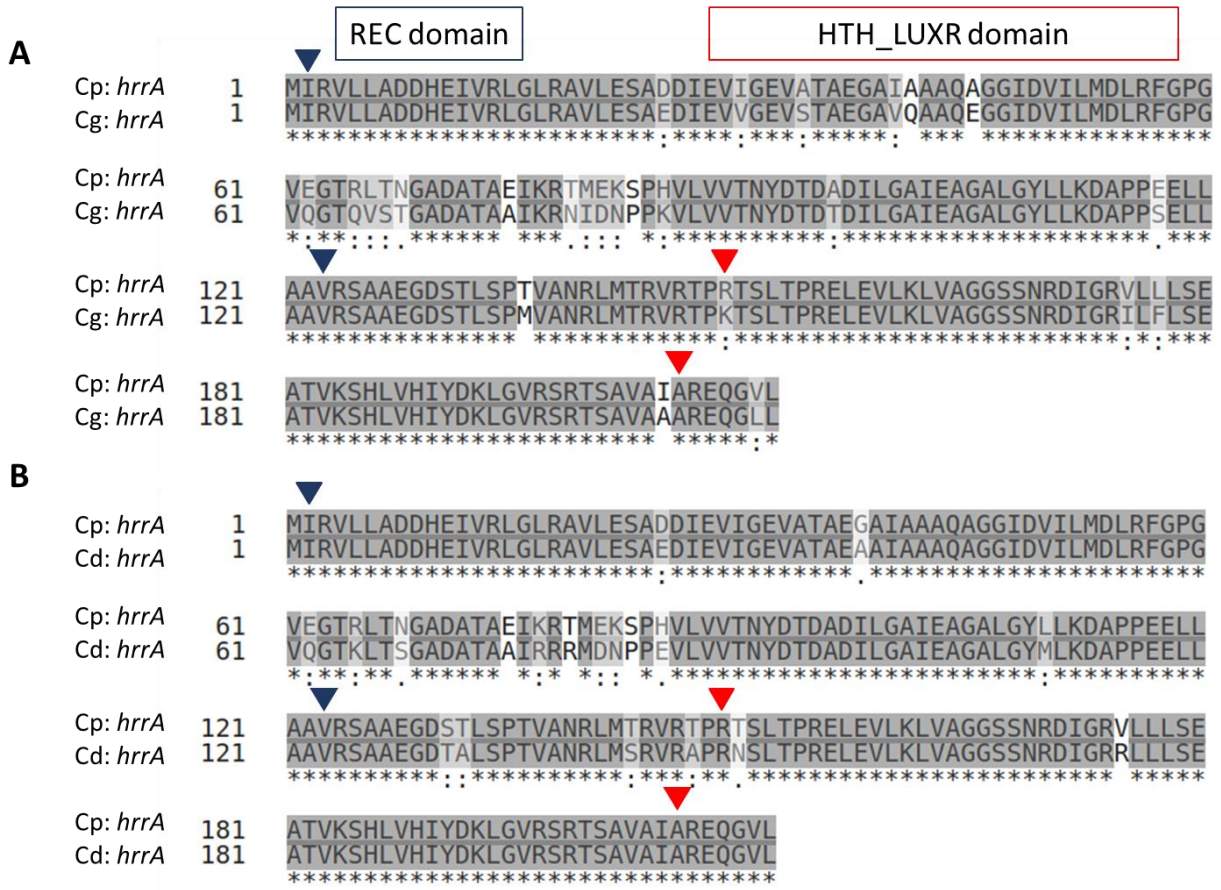
A



B



Additional file 5: Figure S7. Alignment of the amino acid sequences of the HmuT protein between *C. pseudotuberculosis* (Cp), *C. ulcerans* (Cu) and *C. diphtheriae* (Cd). Asterisks indicate sequence identity and similarity is indicated by colons and period. Conserved tyrosine residues (Y233 and Y347) are indicated above the sequence alignment with red arrows (**A**) Cp and Cu alignment with 96.86% amino acid sequence identity; (**B**) Cp and Cd alignment with 80.73% amino acid sequence identity.



Additional file 5: Figure S8. Alignment of the amino acid sequences of the DNA-binding regulator *hrrA* of the two-component regulatory system *hrrSA* between *C. pseudotuberculosis* (Cp), *C. glutamicum* (Cg) and *C. diphtheriae* (Cd). Asterisks indicate sequence identity and similarity is indicate by colons and period. The two domains of *hrrA* are outlined: REC signal receiver regulatory domain is indicated by blue arrow; red arrow indicates HTH_LUXR domain. (**A**) Cp and Cg alignment with 87.73% amino acid sequence identity; (**B**) Cp and Cd alignment with 91.03% amino acid sequence identity.

Additional file 5: Table S9. Genomic Islands (GIs) prediction

			Island location	
	Genomic islands	Genes located within the islands	start	end
T1 strain	Putative Genomic Island 1	CpT1_RS01715-CpT1_10555	359869	365559
	Putative Genomic Island 2	CpT1_RS02765-CpT1_02850	575431	604574
	Putative Genomic Island 3*	CpT1_07140-CpT1_07325	1579138	1619078
	Putative Genomic Island 4	CpT1_08725-CpT1_10755	1913406	1921880
	Putative Genomic Island 5	CpT1_09180-CpT1_09230	2013332	2022451
	Putative Genomic Island 6*	CpT1_09890-CpT1_09925	2193668	2204043
	Putative Genomic Island 7*	CpT1_10290-CpT1_10355	2291646	2305455
1002B strain	Putative Genomic Island 1	Cp1002B_01700-Cp1002B_10545	359878	365568
	Putative Genomic Island 2	Cp1002B_02750-Cp1002B_02835	575437	604583
	Putative Genomic Island 3*	Cp1002B_02990-Cp1002B_03010	637219	642946
	Putative Genomic Island 4*	Cp1002B_03040-Cp1002B_03075	648004	658378
	Putative Genomic Island 5	Cp1002B_10625-Cp1002B_04255	930172	938645
	Putative Genomic Island 6*	Cp1002B_05640-Cp1002B_05830	1232999	1272932
	Putative Genomic Island 7*	Cp1002B_10260-Cp1002B_10325	2289552	2303361

Note: Island prediction was carried out using the T1 and 1002B reference genomes against a closely related nonpathogenic genome (NC_003450). Island location and locus identification of genes within these islands are shown. 7 GIs were predicted in the genome of both strains. *Indicates a putative island harboring differentially expressed genes identified under iron restriction. The genes *htaG-htaF* were identified in the genomic island 4, *Cp3070-Cp_3075* island 6, *Cp_2995* island 7 and *Cp_10265* island 3.

8.2. ANEXO II: Artigo publicado no periódico GENE (ISSN: 0378-1119) - Transcriptome analysis of *Corynebacerium pseudotuberculosis* biovar Equi in two condition of environmental stress (doi: 10.1016/j.gene.2018.08.028).

Anne Cybelle Pinto Gomide[1*§], Izabela Coimbra Ibraim[1*], Jorianne TC Alves[2], Pablo Gomes de Sá[3], Yuri Rafael de Oliveira Silva[2], Mariana Passos Santana[1], Wanderson Marques Silva[4], Edson Luiz Folador[5], Diego C. B. Mariano[6], Thiago Luiz de Paula Castro[7], Silvanira Barbosa[2], Fernanda Alves Dorella[8], Alex F. Carvalho[8], Felipe L. Pereira[8], Carlos A. G. Leal[8], Henrique C. P. Figueiredo[8], Vasco Azevedo[1], Artur Silva[2], Adriana Ribeiro Carneiro Folador[2]

[1]Department of General Biology, Institute of Biological Sciences, Federal University of Minas Gerais -Av. Antônio Carlos, Belo Horizonte 31.270-901, Brazil. Tel: +55(31) 34092873.

[2] Laboratory of DNA Polymorphism, Institute of Biological Sciences, Federal University of Pará - Rua Augusto Corrêa, Belém 66.075-110, Brazil

[3] Federal University Rural of Amazon. Rodovia PA 140, 2428, Tomé-Açu, PA - Brazil

[4] National Institute of Agricultural Technology. Los Reseros y Nicolás Repetto, Hurlingham - 1686, Argentina. Tel: (11) 46211127 Ramal: 109

[5] Biotechnology Center, Federal University of Paraíba. João Pessoa, Brazil

[6] Department of Computer Sciences, Institute of Exact Sciences. Federal University of Minas Gerais -Av. Antônio Carlos, Belo Horizonte 31.270-901, Brazil. Tel: +55(31) 34095810

[7] Departament of Biointeraction, Institute of Health Sciences. Federal University of Bahia – Av. Reitor Miguel Calmon, s/n, Vale do Canela, Bahia, Brazil

[8]AQUACEN - National Reference Laboratory of Aquatic Animal Diseases, Ministry of Fisheries and Aquaculture, Veterinary School, Federal University of Minas Gerais, Belo Horizonte, Brazil

*These authors contributed equally to this work.

[§]Corresponding author: acybelle@gmail.com

Email addresses:

ACPG: acybelle@gmail.com
ICI: izzy.coimbra@gmail.com
JTCA: joriannealves@gmail.com
PHCGS: pablogomesdesa@gmail.com
YROS:ynxs22@gmail.com
MPS: santana.maripassos@gmail.com
WMS: silvamarques@yahoo.com.br
ELF: elf@cbiotec.ufpb.br
DCBM: diegomariano@ufmg.br
TLPC: castrotlp@gmail.com
SB: nirabarbosa@gmail.com
FAD: fernandadorella@gmail.com
AFC: alexficar@gmail.com
FLP: felipe@flpsw.com.br
CAGL: carlosleal@vet.ufmg.br
HCPF: figueiredoh@yahoo.com
VA: vasco@icb.ufmg.br
AS: asilva@ufpa.br
ARCF: adrianarc@ufpa.br

8.2.1. Introduction

Corynebacterium pseudotuberculosis is a Gram-positive, pleomorphic, facultative intracellular pathogenic bacterium belonging to the *Corynebacterium*, *Mycobacterium*, *Nocardia*, and *Rhodococcus* (CMNR) group of the Actinobacteria phylum. It typically has a cell wall consisting of peptidoglycan, arabinogalactan, mycolic acids (BAIRD; FONTAINE, 2007b; DORELLA, FERNANDA ALVES *et al.*, 2006) and a genome with approximately 50% G+C content with an average size of 2,3Mb. This bacterium has two biovars, Ovis and Equi, which are classified based on ability to express the enzyme nitrate reductase; biovar Ovis is negative and biovar Equi is positive for nitrate reduction (RUIZ *et al.*, 2011). In some bacteria the presence of genes related to nitrate respiration confers resistance to stress inside the host, infer an association of nitrate respiration with virulence and pathogenicity (TAN *et al.*, 2010) but in *C. pseudotuberculosis* is not know the role of nitrate respiration.

C. pseudotuberculosis infects a variety of hosts, including camelids, cattle, horses, buffaloes, and humans (ALEMAN *et al.*, 1996; DORELLA, FERNANDA ALVES *et al.*, 2006; PEEL *et al.*, 1997) . Among the infections caused by this pathogen, ulcerative lymphangitis caused by the Equi biovar, which mainly affects horses, results in large losses to agribusinesses (ALEMAN *et al.*, 1996; PEEL *et al.*, 1997). In this disease, *C. pseudotuberculosis* attacks lymph vessels, producing abscesses scattered throughout the body (RADOSTITS *et al.*, 2002) it is thus a microorganism of veterinary significance(ARSENAULT *et al.*, 2003; GUIMARÃES *et al.*, 2011; WILLIAMSON, 2001).

Microorganisms face various adverse conditions when challenging host defense mechanisms during the infectious process, especially within macrophages. These conditions include changes in pH, temperature, osmotic pressure, and nutrient availability as well as the presence of toxic molecules, including nicotinamide adenine dinucleotide phosphate (NADPH), hydroxyl radicals, and hydrogen peroxide (H_2O_2). However, many bacteria are able to adapt to these conditions by triggering modulatory responses in the host cell, thereby promoting their own survival and proliferation. In some pathogenic bacteria, the ability to trigger such modulatory responses is associated with virulence (MEIBOM *et al.*, 2008; SCHUMANN, 2007).

To escape deleterious conditions within the potential host, the activation of genes directly or indirectly related to virulence in a manner that is regulated by the extracytoplasmic function (ECF) sigma factors are essential (KAZMIERCZAK; WIEDMANN; BOOR, 2005). An adaptive response is thus generated in response to various environmental changes. The mutation in the gene encoding sigma factor E in *Mycobacterium tuberculosis* and *Haemophilus influenzae*, for example, demonstrated that possess effective mechanisms for escaping the immune response and establishing bacterial persistence in mice or in human macrophage cells, thereby contributing to intracellular replication and to the regulation of responses to various stresses (CASONATO *et al.*, 2014; CRAIG; NOBBS; HIGH, 2002; MANGANELLI *et al.*, 2001).

The data of *C. pseudotuberculosis* related to the presence and activity of virulence factors, as well as studies on the activation of sigma factors under different conditions, are increasingly being reported (PACHECO *et al.*, 2012; SILVA; SEYFFERT; CIPRANDI; *et al.*, 2013; SILVA; SEYFFERT; SANTOS; *et al.*, 2013). Transcriptome studies on

strain 1002, biovar Ovis, using the *de novo* approach and differential analysis expression using the RNAseq method have inferred possible candidate genes, including *dps, msrB*, and the genes encoding various sigma factors, that are involved in the persistence and survival of the bacterium under unfavorable conditions such as acidic stress, heat shock, and osmotic stress (PINTO, ANNE *et al.*, 2014; PINTO, ANNE CYBELLE *et al.*, 2012). Thus, the use of this technology has contributed in identifying new possible candidate virulence factors.

Among the most studied genes of *C. pseudotuberculosis* are *pld* (phospholipase D) and *fagB* (iron siderophore binding protein) genes, whose functions are essential for the development of lymphangitis. The PLD protein is an exotoxin that aids in the dissemination of the organism from the infection site to the lymph nodes (MCKEAN; DAVIES; MOORE, 2007a; MCNAMARA; BRADLEY; SONGER, 1994; MCNAMARA; CUEVAS; SONGER, 1995), whereas *fagB* encodes an iron acquisition protein that enables the microorganism to obtain this element from its host (AQUINO DE SÁ *et al.*, 2013; BILLINGTON; ESMAY; SONGER; JOST, 2002).

The RNAseq tool has increasingly contributed to transcriptome studies in prokaryotes because of its advantages relative to other expression analysis technologies, making it an attractive technology for expression profiling studies (HAAS *et al.*, 2012; OSMUNDSON; DEWELL; DARST, 2013).

However, the real-time PCR methodology it is still used to validate many of the results obtained, being considered the gold standard for the quantification of differently expressed genes (DERVEAUX; VANDESOMPELE; HELLEMANS, 2010). Although, for many researches, NGS technology is already capable of presenting better and more precise results than real-time PCR(LADETTO *et al.*, 2014; SENDLER; JOHNSON; KRAWETZ, 2011).

In an effort to understand the biological evolution, that is, the adaptation of *C. pseudotuberculosis* to constantly changing environments, the present study aimed to examine the transcriptomics of *C. pseudotuberculosis* strain 258 biovar Equi under conditions faced by the bacterium in host cells, including acidic stress and osmotic stress, through RNA-seq. The evaluation of the transcriptional profile of this organism has

provided information that will contribute to future studies aimed at treating, reducing or eradicating ulcerative lymphangitis in animals.

8.2.2. Materials and Methods

Genomics

Optical map of *C. pseudotuberculosis* 258

Complete mapping of the *C. pseudotuberculosis* 258 strain was performed using the Argus technology of OpGen (http://opgen.com/genome-mapping-products-services/softwares/mapsolver), which offers high resolution. Whole-genome restriction maps were sorted based on a single DNA molecule. To design the restriction map with DNA fragments ranging from 200 kb to 2.5 Mbp in length, a solid phase was used to capture the DNA and the enzyme *Kpnl* was used in the digestion step. The final analysis was performed using a fluorescence microscope to convert the optical map into digital data. The software MapSolver (http://opgen.com/genome-mapping-products-services/argus-system) was used to visualize and map the data for comparison with the genomic sequence.

***C. pseudotuberculosis* 258 genome assembly and annotation**

The *Corynebacterium pseudotuberculosis* 258 genome was re-sequenced in the platform Ion Torrent PGM™ System, using the fragment library and chip 318. The analysis of reads quality was performed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The reads were subjected to *de novo* assembly using Mira 3.9 (CHEVREUX; WETTER; SUHAI, [S.d.]). A modified version of the software CONTIGuator 2.7 (GALARDINI *et al.*, 2011) was used for the orientation and sorting of contigs, which used *C. pseudotuberculosis* 258 version 1 sequenced by the platform SOLiD V3 as the reference genome. The scaffolding was performed by MapSolver™ (http://opgen.com/genomic-services/softwares/mapsolver), in addition, both software SIMBA (http://lgcm.icb.ufmg.br/simba) and CLC Workbench 7 (http://www.clcbio.com) were used for the gap closing stage.

The second version of the *C. pseudotuberculosis* 258 genome was automatically annotated with the software Rapid Annotation using Subsystem Technology (RAST) (AZIZ *et al.*, 2008). Error correction in homopolymer regions was performed using the software Artemis (RUTHERFORD *et al.*, 2000), and the UniProt

(http://www.uniprot.org) database was used for manual correction of automatic annotation.

Transcriptomics

Bacterial cultures

Corynebacterium pseudotuberculosis biovar Equi strain 258, isolated from a horse in Belgium, was grown in plates with Brain Heart Infusion Broth (BHI) consisting of 200 g calf brain infusion, 250 g beef and heart infusion, 10 g protease peptone, 2 g dextrose, 5 g sodium chloride and 2.5 g disodium phosphate; pH 7.4 ± 0.2 at 25°C at room temperature (RT). For the pre-inoculum, one colony was inoculated into 20 ml of liquid BHI media in a Falcon tube supplemented with Tween® 80 (0.05%) for a 24-hour period at 37ºC in a shaker at 160 rpm. Following this period, a 1/100 dilution was performed to prepare the inoculum in fresh BHI media supplemented with 80% Tween® (0.05%), followed by incubation at 37°C in a shaker at 160 rpm. The growth of the bacterial culture was monitored for 14 hours (see Additional file 1).

Application of stress conditions

The experiment was carried out at the beginning exponential phase so that we could observe, at the onset of replication, the expression profile that is able to model a rapid and adaptive direct response, seeking survival. When the inoculated culture entered an exponential phase (A_{600} = 0.2), it was divided into two Falcon tubes (20 ml each), which were centrifuged for three minutes at 8000 rpm at 37ºC. The supernatant was discarded, and the pellet was resuspended in BHI supplemented with stressors specific to each condition. To simulate osmotic stress, 2 M NaCl was added to the growth medium; for acid stress, hydrochloric acid was added until the pH stabilized at 5.0. A control was prepared by resuspending the pellet in BHI medium under physiological conditions. The tubes were incubated at 37°C in a shaker at 160 rpm for 15 min. An aliquot taken from each tube was then serially diluted to 10^{-1} to 10^{-6}. Dilutions of 10^{-4}, 10^{-5} and 10^{-6} were plated on BHI agar, and the plates were incubated at 37°C for 48 hours followed by a determination (in duplicate) of cell viability and colony counting. The remainder of each sample was centrifuged at room temperature for three minutes at 8000 rpm, and the pellet was resuspended in 2 ml RNAlater® (Ambion, USA) according to the manufacturer's instructions.

RNA extraction

The recovered cells were subjected to total RNA extraction using the ChargeSwitch® Total RNA Cell kit (Invitrogen) according to the manufacturer's protocol with some modifications. For the cell lysis step, glass microbeads 1 mm in diameter (Bertin Technologies) were added to 2-ml tubes containing the samples, and the tubes were homogenized using a Precellys 24 homogenizer at 6,500 rpm for two 15-second cycles with a 30-second interval between cycles. The samples were then centrifuged for one minute, and the supernatant was transferred to a new 2-ml tube and incubated in a dry bath at 60ºC for 15 minutes. DNase was added to avoid any contamination with genomic DNA. The total RNA extracted was eluted in 100 µL of RNase-free milli-Q water and quantified using a Qubit® 2.0 fluorometer (Invitrogen, USA).

rRNA depletion for mRNA enrichment

mRNA enrichment was performed through rRNA depletion using the Ribominus™ Transcriptome Isolation kit for yeast and bacteria (Invitrogen, USA), according to the manufacturer's instructions.

SOLiD™ V3 sequencing

From the enriched mRNA sample, the cDNA library was constructed, using the SOLiD™ V3 Total RNA-Seq kit (Applied Biosystems, USA) according to the manufacturer's instructions, and quantified using a Qubit® 2.0 fluorometer (Invitrogen). In the library construction process, the RNA was fragmented through the RNase III enzyme, end adapters were added, and the sample was amplified via reverse transcriptase, following the protocol provided with the SOLiD™ Total RNA Seq kit (Life Technologies™, CA). Fragments in the 150 to 250 base pair range were, then, selected via polyacrylamide gel electrophoresis, 6% of the cDNAs amplified by PCR. They were purified using the PureLink™ PCR Micro kit (Invitrogen, USA) and quantified using a Qubit® 2.0 fluorometer (Invitrogen, USA). Samples were confirmed through electrophoresis on 2% agarose gel.

Through the Applied Biosystems SOLiD ™ 3 Plus System Templated Bead Preparation Guide, it was possible to develop the PCR in the emulsion, where the amplification is performed using primers complementary to the microbead adapters. After this step, microbeads containing sequences of interest were captured and deposited on slides to

perform the sequencing on the SOLiD™ V3 platform, according to the manufacturer's recommendations.

Transcriptomic analysis

Two files were obtained following sequencing: one in csfasta format and another in QVqual format; these files contained the reads generated in color space and the quality of the bases of said reads (quality value, QV), respectively. The software FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to assess the quality value of bases using Phred values equal to or greater than 20.

Following the reads quality analysis, transcripts from each library were aligned to the *C. pseudotuberculosis* 258 genome using the software TopHat v2.0.4, which uses the aligner Bowtie v2.0.0.5 (Sasindran et al., 2007a). The bam/sam files generated by TopHat v2.0.4 for each experimental condition were input into the program Cuffdiff v2.0.2, part of the software suite Cufflinks (Trapnell et al., 2010), which makes it possible to obtain expression values in FPKM (Fragments Per Kilobase Of Exon Per Million Fragments Mapped). The results generated using Cuffdiff v2.0.2 were then subjected to statistical analysis using the one-class Significance Analysis of Microarrays (SAM) method, which tests whether the mean expression values differ from zero (Tusher et al., 2001) and identifies genes with significant differential expression.

For the selection of differentially expressed genes, genes showing at least a two-fold change in expression relative to control were considered. The fold change was calculated as the ratio of FPKM values in the stress and control conditions. The functional analysis of these genes was performed using the software Blast2GO (http://www.blast2 go.com), which used a gene ontology database.

RT-qPCR assay.

The total bacterial RNA was reversed transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). cDNA syntheses were carried out in the ATC 401 (Nyx Technik) thermocycler and samples were storaged in the -80C for posterior quantitative real-time analysis qPCR.

Quantitative analysis was performed using the 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific) and primers with PCR amplification efficiency between 90-100% (Table 1) Analyses were performed in triplicates for each of the condition and genes evaluated. PCR was carried out in 96-wells optical plates (Thermo Fisher Scientific) and

cDNA real-time amplification results were analyzed with the SDS (Thermo Fisher Scientific) software. All experiments were done following recommendation (Bustin et al., 2009).

8.2.3. Results and Discussion

Genomics

***C. pseudotuberculosis* 258 genome refinement**

The bacterium *C. pseudotuberculosis* 258 was subjected to optical mapping analysis using the MapSolver™ software.

Comparison of the results of this analysis with the first published version of the genome (accession number CP003540.1) showed that the whole-genome map present a genome size bigger, which can indicate the occurrence of missing regions (see Additional file 2A). The quality analysis of the data showed that the reads on average had Phred qualities greater than 20, indicating no need for quality filter application. *De novo* assembly performed using the software Mira (CHEVREUX; WETTER; SUHAI, [S.d.]) showed 41 contigs, 19.59x average coverage, and an N50 value of 93,650. The scaffolding of the contigs was performed by MapSolver™ software (see Additional file 2B) and the gap closing by SIMBA and CLC Workbench 7. After completion of the assembly, the genome size was 2,369.817 bp, an increase of 55,413 bp over the original version (see Additional file 2C).

The annotation of the new sequence showed 43 new coding sequences (CDS), and the number of pseudogenes decreased from 46 to 33. Additional file 2D shows a 17-kb region identified using the Artemis Comparison Tool (ACT) (CARVER *et al.*, 2005) wherein genes composing the operon *narKGHJI*, which is responsible for nitrate reduction, were identified (NISHIMURA *et al.*, 2011). This region was missing in the first version of the sequence.

Transcriptomics Analysis of differentially expressed genes

To evaluate the gene expression profile of *C. pseudotuberculosis* 258, stress-causing agents were added to the bacterial growth medium during the early exponential phase of growth ($A_{600nm} = 0.2$). Cell viability was assessed through the colony-forming unit assay. The results showed a 36% decrease in replication under acid stress, a 16% decrease under osmotic stress (see Additional file 3). This decrease in replication occurs naturally due to the microorganism's adaptation to environmental conditions (JOZEFCZUK *et al.*, 2010).

In an analysis of the transcript mapping results with respect to the reference genome, the control condition showed the greatest number of uniquely aligned transcripts, followed by the acid and osmotic stress conditions (Table 2).

The software Cufflinks was used to obtain FPKM (Fragments Per Kilobase of Exon Per Million Fragments Mapped) values (TRAPNELL *et al.*, 2010) for each gene tested. Of the 2131 coding sequences (CDS), 2019 (94.7%) were expressed in the control condition, and 1906 (89.4%), 1766 (82.8%) were expressed under the acid and osmotic respectively.

The *p-value* < 0.001 was used to select genes that exhibited significant differential expression. The selection of these genes was based on fold change such that only genes with an expression ratio at least twice that of the control were selected for analysis. Thus, genes considered induced showed fold changes ≥ 2, whereas repressed genes showed fold changes ≤ 0.5.

Under conditions of acid stress, 813 genes were considered differentially expressed. A total of 480 genes had fold-change cutoff ≥ 2 or ≤ 0.5; of these, 272 were induced and 208 were repressed. Under osmotic stress conditions, 501 genes were differentially expressed, including 115 induced and 194 repressed genes (309 genes total) with the fold change ≥ 2 or ≤ 0.5 (Fig. 1). The genes scores within the cutting set were considered, including those present in two conditions at the same time.

Biological processes

Gene ontology (GO) analyses showed that 22 biological processes were involved in the response to acid stress and in the osmotic stress, 21 processes were involved (Gene Ontology level 3; Fig. 2 and 3). Among the processes related to bacterial defense mechanisms for survival under adverse conditions, we highlight the following: “response to stress”, “regulation of biological process”, and “response to chemical stimulus”. When analyzing the genes involved in each process, the differential expression of each individual gene in more than one condition and even in more than one process was observed, suggesting that some genes have different functions in different unfavorable environments (see Additional file 4: Table S1).

Induced genes involved in the "stress-response process"

The stress-response process is related to conditions of cellular imbalance, for example, the environment faced by bacteria when infecting hosts. It is noteworthy that during phagocytosis, the internal pH of macrophages decreases (pH = 5.0), which is harmful to microorganisms. In this environment, the tendency to generate reactive oxygen species is even greater (FARR; KOGOMA, 1991). Thus, bacterial molecules undergo redox reactions with ions and molecules present in the environment to survive, thus achieving internal homeostasis (FOLLMANN *et al.*, 2009). When cells face adverse conditions of this type, proteins undergo changes that impair their function and require protection. We highlight the following genes based on this profile: *msrA, dps* in the acid stress conditions. The variation in expression of these genes is notable, showing the environmental effect on the transcription and regulation of each gene (see Additional file 4: Supplementary Table S1). No one was shared between the conditions. Only 3 genes were induced in the osmotic stress and 2 out of these genes were found to encode hypothetical proteins. The repair protein methionine sulfoxide reductase is encoded by the genes *msrA* and *msrB*. These genes share no homology and display the same enzymatic activity, albeit operating on different substrates (SASINDRAN; SAIKOLAPPAN; DHANDAYUTHAPANI, 2007). Only *msrA*, which showed an 8.7-fold change under acid stress (Fig. 4) was considered induced in strain 258 (Equi), whereas *msrB* was induced only under acid conditions in strain 1002 (Ovis) (PINTO, ANNE *et al.*, 2014). Previous studies have shown that the role of MsrA is more relevant to virulence than MsrB (SASINDRAN; SAIKOLAPPAN; DHANDAYUTHAPANI, 2007). However, strain 258 is considered more virulent than 1002, which has a low virulence profile (WMS, 2013 - unpublished observations). Studies performed in *Mycobacterium smegmatis* and *Leishmania major* involving *msrA* inactivation by gene recombination showed that the mutant strains underwent significantly reduced replication within macrophages and were more sensitive to hydrogen peroxides (H_2O_2) than the wild-type strains (DOUGLAS *et al.*, 2004; SANSOM *et al.*, 2013). The *msrA* gene has a key role in bacterial adhesion to host cells; in *Helicobacter pylori*, the absence of this gene resulted in reduced virulence in studies conducted on animal models (ALAMURI; MAIER, 2004). These results provide strong evidence for the importance of this gene in contributing to the resilience of bacteria in harsh environments and suggest that induction of this gene in the Equi biovar may exert a beneficial activity within macrophages.

The *dps* gene, which showed a 10.2 fold change in expression under acid stress conditions (Fig. 4), is related to DNA binding and iron uptake from the medium. Its gene product, which possesses ferroxidase activity, reduces the amount of toxic peroxides present in the medium through iron oxidation catalysis (CALHOUN; KWON, 2011). Its function contributes to bacterial protection against acid, oxidative and heat stress, UV exposure, gamma radiation and iron and copper toxicity during the stationary phase (CALHOUN; KWON, 2011; NAIR; FINKEL, 2004). A study conducted on *Escherichia coli* mutants in which the *dps* gene had been inactivated showed a significant reduction in the survival rate of the mutants in the logarithmic and stationary phase of bacterial replication compared to wild-type *E. coli* when subjected to an acid environment (pH = 1.8) (CHOI; BAUMLER; KASPAR, 2000). The *dps* gene was also induced under acid stress in *C. pseudotuberculosis* 1002 (PINTO, ANNE *et al.*, 2014), indicating that it is an important candidate for studies related to fighting caseous lymphadenitis and ulcerative lymphangitis.

Induced genes involved in the regulation of biological processes

The regulation of biological processes forms the basis of living beings' ability to maintain Equilibrium of all cell processes under various conditions. This regulation may occur in a number of ways, including the regulation of transcription, which may be confirmed by the presence of regulators and sigma factors in this process.

Transcriptional regulators are key players in the modulation of adaptive responses in the microorganism. Nine genes showed changes in expression under osmotic stress conditions and 21 showed changes under acid stress conditions. Only two genes were shared among the conditions (see Additional file 4: Supplementary Table S1). We highlight the following genes in this process: *tetR2, glmU* (Fig. 4). The *lysR1 and sigB, C e E* (Fig. 4) genes were present only in the acid condition and we highlight due to the fold change, but only *lysR1* gene was related in this process. The sigma factor genes were listed in the item about sigma factors.

The *tetR2* gene, which showed an approximately a nine-fold change under osmotic stress conditions and an 18 fold change under acid stress conditions (Fig. 4) encodes a transcriptional regulator of the TetR family. This family is involved in several processes, including the transcriptional regulation of enzymes in catabolic pathways, the

biosynthesis of antibiotics, drug resistance, differentiation and pathogenicity (RAMOS *et al.*, 2005). In a study of the plant bacterium *Acidovorax avenae*, deletion of the *tetR* gene showed the importance of this regulator for resistance to oxidative stress, cell replication, and biofilm formation, all of which were reduced compared to the wild-type strain. However, direct regulation of virulence genes by TetR was not observed (LIU *et al.*, 2014). This result suggests that the set of genes regulated by TetR in *C. pseudotuberculosis 258* is involved in responses to various stresses and that these genes may contribute to the resilience of these bacteria in adverse environments.

The gene *glmU*, which showed, a three-fold change under osmotic stress and a 2.8–fold change under acid stress (Fig. 4), encodes a bifunctional enzyme with glucosamine-1-phosphate acetyl transferase and N-acetylglucosamine-1-phosphate uridylyl transferase activity. This enzyme catalyzes the formation of UDP-GlcNAc, a key precursor in the synthesis of bacterial peptidoglycans and lipopolysaccharides, which confer strength and stiffness to the cell wall (DOIG *et al.*, 2014). To assess the importance of this enzyme in stress responses in *Mycobacterium smegmatis*, the *glmU* gene was deleted. The results showed that this enzyme is essential for bacterial replication; the cells showed modified morphology and/or were lysed. Thus, this enzyme represents a key target for the development of anti-mycobacterial drugs (ZHANG *et al.*, 2008).

Cell-surface lipids have been described as factors that contribute to bacterial pathogenicity in *C. pseudotuberculosis*. The presence of lipids associated with the cell wall of this bacterium provides resistance to digestion by cellular enzymes, enabling the microorganism to persist as a facultative intracellular parasite (DORELLA, FERNANDA ALVES *et al.*, 2006). According to the aforementioned reports, the induction of *glmU* in *C. pseudotuberculosis* 258 under all stress conditions simulated *in vitro* suggests that this gene is highly important for bacterial resistance to adverse conditions and may contribute to its persistence in hosts.

The gene *lysR1*, a transcriptional regulator of the LysR family encoding a transcriptional activator protein, was overexpressed only under acid stress conditions (6.7-fold change; Fig. 4). Members of this gene family affect the regulation of various virulence processes, modulate responses to various stresses and act as activators or repressors of a variety of genes (MCCARTHY *et al.*, 2014; REEN *et al.*, 2013). This gene was also overexpressed in *C. pseudotuberculosis* 1002 only under acid stress conditions (seven-fold change)

(PINTO, ANNE *et al.*, 2014). These results suggest that *lysR1* plays a modulatory role to produce a satisfactory response in both strains. Moreover, these global regulators have been the subject of many studies focused on virulence, diagnosis, and therapy, and the *lysR1 gene* may be a candidate for further studies on *C. pseudotuberculosis* seeking effective solutions against both ulcerative lymphangitis and caseous lymphadenitis.

Genes involved in responses to chemical stimuli

Responses to chemical stimuli refers to changes in the organism resulting from chemical stimulation. Only three genes were classified in this process. The genes *trxB*, *norM* and *uppP* (Fig. 4) were notable because of their expression values, which were higher under the conditions tested (see Additional file 4: Supplementary Table S1).

The *trxB* gene, which showed a 7.2-fold change under acid stress condition encodes a thioredoxin reductase that, together with thioredoxin (trxA), forms part of the thioredoxin system (Trx) (ASANO; DAVIES, 1998). Studies performed in *M. tuberculosis* showed that this bacterium is able to resist oxidative stress within phagocytes through thioredoxins, suggesting that the elimination of free radicals produced by mononuclear phagocytes during infection is a mechanism of resistance to intracellular death (SHINNICK; KING; QUINN, 1995). In *Neisseria gonorrhoeae*, a human pathogen, the *trxB* mutant showed greater sensitivity than a wild-type strain to oxidative/nitrosative environments, and the protein TrxB was shown to be require d for biofilm formation (POTTER *et al.*, 2009). These results emphasize the importance of the *trxB* gene in *C. pseudotuberculosis* 258 and suggest that it may contribute to the resistance of the bacterium to adverse environments.

The *norM* gene, which showed a four-fold change under osmotic stress conditions and a three-fold change under acid stress conditions (Fig. 4), encodes a multidrug resistance protein belonging to the ATP-binding cassette (ABC) transporter family. This protein provides effective resistance to antibiotics because it encodes a multidrug efflux pump, and it is considered a potential virulence factor in some bacterial species (BRAIBANT; GUILLOTEAU; ZYGMUNT, 2002; MORITA *et al.*, 1998). The ABC transporters promote the extrusion of toxic molecules and drugs via a proton/cation electrochemical gradient through the membrane, thus contributing to intracellular homeostasis (VANNI *et al.*, 2012). In *C. pseudotuberculosis* 1002 (Ovis), *norM* may also be induced under

osmotic stress conditions with a significant fold change (3.2) (PINTO, ANNE *et al.*, 2014), favoring its potential use in the control of a variety of diseases in different hosts.

The *uppP* gene encodes an Undecaprenyl-diphosphatase enzyme and presented a little over 2 fold change in the osmotic stress in relation to the control. In bacterias, such as *E.coli*, *Burkholderia sp*., *Bacillus subtilis*, amongst others, it is related to the biosynthesizes of different cell wall components and membrane proteins, thus indicating that this gene plays a major role in bacterial integrity and suggesting that its absence could lead to an increased in bacterial susceptibility to abiotic stress. Consequently, reducing colonization and diminishing bacterial resistance against the hosts immune response (KIM *et al.*, 2013; ZHAO *et al.*, 2016). These genes, as well as the *glmU* discussed above, plays a pivotal role in the cell wall constitution, colonization resistance, replication and installation of the pathogenic process. Thus, becoming an important antibacterial target.

Expression of sigma factor-encoding genes

The genome of *C. pseudotuberculosis* strain 258, as identified by (PINTO, ANNE *et al.*, 2014) in strain 1002, contains eight sigma factor-encoding genes: SigA, which is primary and essential; SigB, which is alternative and non-essential; and SigC, SigH, SigD, SigE, SigK, and SigM, which encode alternative and extracytoplasmic factors.

Acid stress was the only condition that resulted in induction of sigma factor-encoding genes to the extent established by the expression cutoff (Fig. 5). The induced genes included *sigB (*fold change = 7.0), *sigC* (fold change = 4.2) and *sigE* (2.0), which were included in the biological process of regulation (see Additional file 5: Supplementary Table S2).

Sigma B is an alternative factor involved in the responses of various Gram-positive bacteria to stress, including heat, acid and alcohol stress. An *L. monocytogenes* strain with deletion of this sigma factor showed no reduction in viability under heat stress, although it showed nearly ten thousand times more susceptibility to lethal acid stress (pH = 2.5) (FERREIRA; O'BYRNE; BOOR, 2001). In *C. pseudotuberculosis* 258, the increased expression of sigma B compared to the control suggests a strong effect of this factor in

the regulation of genes require d for bacterial adaptation to acid environments, including the macrophage environment in hosts.

The gene *sigE*, which was also induced in response to acid stress, was essential for the survival of *M. tuberculosis* in macrophages and is apparently involved in the bacterium's resistance to oxidative agents. The same study also showed that the sigma E factor may affect the expression of the alternative sigma factor, SigB (MANGANELLI *et al.*, 2001). These genes and their regulons may form part of a gene catalog to be tested in studies involved in the development of therapies, diagnoses or vaccines against caseous lymphadenitis and ulcerative lymphangitis because *sigB* and *sigE* were also found to be significantly induced, with fold changes greater than two, in *C. pseudotuberculosis* 1002 (PINTO, ANNE *et al.*, 2014).

The sigma C factor is associated with virulence in pathogenic bacteria of the genus *Mycobacterium*. It is conserved in pathogenic species and absent in non-pathogenic species and is one of only two sigma factors with extracytoplasmic function that have been found in *M. leprae* (RODRIGUE *et al.*, 2006). In *M. tuberculosis*, sigma C was deemed non-essential for bacterial survival inside macrophages, although deletion of its gene decreased virulence in this species (SUN *et al.*, 2004). In *C. pseudotuberculosis* 258, sigma C was shown to be involved in adaptive responses because it was considered differentially expressed and was induced at the beginning of the exponential phase of growth in acid medium.

Expression of genes present in pathogenicity islands

Virulence genes form the basis of certain adaptation mechanisms and contribute to the survival of bacteria during the infection of hosts. Virulence genes are usually acquired by horizontal transfer, are expressed when the bacterium comes into contact with the host, and are located in regions called pathogenicity islands (GAL-MOR; FINLAY, 2006; HARE; HUEFFER, 2014).

A total of 16 pathogenicity islands were predicted in *C. pseudotuberculosis* 258, totaling 265 genes (SOARES *et al.*, 2012, 2013). Of these, 78 genes were considered differentially expressed under at least one condition (see Additional file 6: Supplementary Table S3).

Fig. 6 shows a heat map in which color change is used to represent the expression levels of individual genes under each stress condition. The expression levels of genes whose protein products are well characterized in the literature, including *katA*, *pld*, *srtA*, and *clpB*, were assessed.

The *katA* gene (6.4-fold change), induced under acid conditions (Fig. 4), encodes a catalase that acts as an antioxidant by converting hydrogen peroxide into water and molecular oxygen (GIORGIO *et al.*, 2007). KatA is a stable enzyme that is typically highly resistant to several proteases and plays a key role in the adaptation to H_2O_2 stress and in virulence (BANDYOPADHYAY; STEINMAN, 2000; LEE *et al.*, 2005). The *katA* gene is expressed at all stages of bacterial growth but especially in the stationary phase (SHIN; CHOI; CHO, 2008). Several studies have already been conducted in attempts to understand the role of this enzyme in the mechanism of virulence of pathogenic bacteria.

In a study on biofilms formed by *Pseudomonas aeruginosa*, *katA* conferred resistance to various stress conditions and antibiotics (SHIN; CHOI; CHO, 2008). Biofilm formation is related to chronic infection and to resistance to antibiotics (COSTERTON; STEWART; GREENBERG, 1999); it has been reported in *C. pseudotuberculosis*, *C. renale* and *C. diphtheriae* (GOMES *et al.*, 2009; OLSON *et al.*, 2002). *C. pseudotuberculosis* bacteria that produced biofilm were resistant to seven antibiotics, including tetracycline and ampicillin (OLSON *et al.*, 2002). The increased synthesis of polysaccharides lining the biofilm contributes to protection against external agents as well as to antibiotic resistance and bacterial stability (COSTERTON; STEWART; GREENBERG, 1999; SMITH, ANTHONY W., 2005).

The gene *pld*, which encodes phospholipase D, was induced only under acid conditions (2.4-fold change; Fig. 4). Phospholipase D is able to hydrolyze lysophosphatidyl choline and sphingomyelin, thereby promoting diffusion of the invading bacterium by damaging endothelial cells and lymphatic and blood vessels (MCKEAN; DAVIES; MOORE, 2007b). In a study of *C. pseudotuberculosis* in which the *pld* gene was inactivated, primary infection was reduced, reinforcing the need for the protein encoded by this gene for the onset of disease and for bacterial persistence in the host (MCKEAN; DAVIES; MOORE, 2007a). Another study evaluated the control of *pld* gene expression under various conditions, concluding that the expression of this virulence factor is reduced when *C. pseudotuberculosis* is exposed to heat shock (43ºC), while its expression is

significantly increased when bacteria infect macrophages. It was shown that in cultures with low cell density, including optical density (OD) = 0.1, the *pld* expression is low and that it may be even lower under heat stress conditions (MCKEAN; DAVIES; MOORE, 2007a, b). These studies corroborate our results in *C. pseudotuberculosis* 258.

The *srtA* gene was induced under acid (7.0-fold change) (Fig. 4). *SrtA* encodes a group of enzymes called sortase that is found in all gram-positive bacteria and is involved in the ability to cause various diseases (PATERSON; MITCHELL, 2006). Sortase A is a cell surface protein that is bound to the cell wall. It modifies other cell surface proteins by cleaving the C-terminus region, thus permitting the binding of proteins that mediate bacterial adhesion.

The *srtA* gene is also considered a housekeeping gene. Its inactivation reduced prevalence during infection, as shown in a study on *Listeria monocytogenes* (BIERNE *et al.*, 2002). The *strA* gene was also shown to be highly important to the pathogenicity of *Streptococcus pneumoniae*, and its role in protein adhesion to the cell surface was demonstrated (PATERSON; MITCHELL, 2006). These results confirm that sortase plays a role in bacterial virulence and is a potential therapeutic target.

The *clpB* gene, which showed a 10.4-fold change under acid stress conditions (Fig. 4), encodes an ATP-dependent chaperone belonging to the Clp/HSP100 gene family and contributes to the biological process of the stress response (see Additional file 4: Supplementary Table S1). *ClpB* is conserved among bacterial species and is important in other chaperone systems, acting in the solubilization and refolding of aggregated proteins (KANNAN *et al.*, 2008; MOTOHASHI *et al.*, 1999; ZOLKIEWSKI, 1999), the disaggregation and degradation of proteins damaged by stress, and especially in helping bacterial cells survive high temperatures by increasing the synthesis of heat shock proteins (HSPs) (NEIDHARDT, 1996; SHIBER; RAVID, 2014). *ClpB* is require d for cellular tolerance of various stresses, including heat, osmotic, ethanol, and acid stress (ISHIKAWA *et al.*, 2010; SQUIRES *et al.*, 1991), according to (KANNAN *et al.*, 2008; LOURDAULT *et al.*, 2011).

A study conducted in *Listeria monocytogenes* using a *clpB* mutant strain showed a 100-fold decrease in the virulence of this strain compared to the wild-type strain (CHASTANET *et al.*, 2004). Therefore, *clpB* may play a role in *C. pseudotuberculosis* 258 resistance during the infectious process inside host macrophage cells.

Distribution of differentially expressed genes in stimulons

The Venn diagram (Fig. 7) shows the distribution of the number of genes that were considered differentially expressed (showing fold changes ≥ 2 or ≤ 0.5) in each stimulon. Under acid stress conditions, more genes were involved in an adaptive response seeking persistence of the bacteria in this environment. Repression of a greater number of genes related to energy metabolism, including aldehyde dehydrogenase, *malK, and cydA*, was also observed (see Additional file 7: Supplementary Table S4). This may explain the results of the colony-forming-unit experiment, in which a smaller number of viable cells were observed (see Additional file 3). This survival strategy reduces bacterial replication to enhance adaptation to the environment (acclimatization), showing that the organism has the ability to sense and adjust to distinct physiological conditions (JOZEFCZUK *et al.*, 2010) and to invest in the expression of specific genes to persist under these conditions.

In the stimulons analyzed, the number of hypothetical protein coding sequences was high and exhibited significant fold change and many presents into pathogenicity islands. This suggests that the proteins encoded by these genes are important in the modulation of rapid and adaptive responses that contribute to bacterial persistence under a variety of stress conditions and probably involved with virulence. It also demonstrates the lack of information on this microorganism, similar to strain 1002 (PINTO, ANNE *et al.*, 2014). A total of 38 genes were induced and shared among the conditions; of these, 42.10% encoded hypothetical proteins (see Additional file 8: Supplementary Table S5). Within islands, 16 genes encoded hypothetical proteins were identified under acid stress, 5 under osmotic stress. These findings indicate the need for further studies to elucidate the roles of these proteins because they may have activities related to the virulence and pathogenicity of *C. pseudotuberculosis* may thus represent important therapeutic targets for the development of vaccines against both the Ovis and Equi biovars of *C. pseudotuberculosis*.

A full list showing the expression profile of all genes in each stimulon may be found in the supplementary material (osmotic medium, see Additional file 9: Supplementary Table S6; acid medium, see Additional file 10: Supplementary Table S7).

RT-qPCR Analysis

Although the RNA-seq technology is considered by many authors as the gold standard for gene expression, many still believe that the gold standard is quantitative PCR, which it will not be replaced by RNA-seq, but it will be used to complement and to obtain more precise results for gene expression (COSTA, CARLOTA, [S.d.]). Therefore, to validate the gene expression profile obtained in this study, quantitative PCR was done in triplicates mimicking the same growth and stress conditions as the ones used for the transcriptomic RNA-seq profiling. In other to improve accuracy, it is essential to use a reference endogenous gene to normalize the quantitative assay. (CARVALHO *et al.*, 2014) suggest that, for *C.pseudotuberculosis*, the *dnaG* gene is one of the best genes to be used as an endogenous control in real-timePCRs assays. After analyzing the genes obtained in the RNA-seq transcriptome, we selected a few genes to be validated by quantitative PCR. The target genes and the primers sequences selected are presented in table 1. In the acid stress condition, with the exception of the *msrA* which was repressed and the *glmU* which did not present a significant gene expression (Fig. 8A) in the quantitative PCR assay, almost every selected gene was confirmed and presented the same expression profile as the one detected by the SOLiD^{TM} platform. In the osmotic stress condition, all the genes were confirmed and presented the same expression profile as the one detected by the SOLiD^{TM} platform (Fig. 8B).

These results suggest that these genes could be important candidates in the development of vaccine targets, diagnoses or therapy in the hopes to eradicate the diseases caused by this bacterium.

8.2.4. Acknowledgements

The authors thank the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and the CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the financial support provided to the authors. The authors also thank the Rede Paraense de Genômica e Proteômica for supporting the development of this work and the Research Support Fundação de Amparo a Pesquisa do Estado do Pará and Universidade Federal de Minas Gerais for funding the study.

8.2.5. Additional files

Additional file 1: **Growth curve of *Corynebacterium pseudotuberculosis* strain 258**. The graph shows the optical density at 600 nm (A600nm) of the culture grown under control conditions for 14 hours. The squares indicate the absorbance values obtained at each time point. Collection of cells for application of stress conditions was performed at OD600nm = 0.2, 2.5 hours after the first measurement.

Additional file 2: *Corynebacterium pseudotuberculosis* 258 genome assembly analysis.
Corynebacterium pseudotuberculosis 258 genome assembly was performed using optical approach whole genome mapping (WGM). Blank regions in the MapSolver software correspond to regions with no alignment; alignments are shaded in blue. (A) Genome sequence of *C. pseudotuberculosis* 258 obtained by sequencing a fragment library with the next-generation genome sequencer SOLiD V3 plus (2,314,404 bp); the orientation of the assembly when compared to the optical map was in the correct order. (B) Scaffolding process of contigs obtained by sequencing simple fragments in the platform Ion PGM™ System for the genome sequence of *C. pseudotuberculosis* 258. (C) Comparison between the Whole-genome mapping with the new genome assembly obtained by Ion PGM™ sequencing. The Fig. shows that the new genome assembly has a length close to the indicated by WGM. In addition, we point out that in comparison with the first published version of the genome, the new assembly of C. *pseudotuberculosis* 258 present an increase of 55,413 bp in the genome length. (D) Synteny graph based on data obtained using ACT between *Corynebacterium pseudotuberculosis* 258 (SOLiD data) (above) and *Corynebacterium pseudotuberculosis* data Ion PGM™ System (below), which shows a 17,000-bp insertion region. Key genes of the *narKGHJI* cluster are mapped to this region.

Additional file 3 - Number of viable cells under each condition tested.
Number of colonies obtained in the colony-forming unit experiment (viability). Blue: control condition. Green: acid stress, in which a 36% reduction compared to control is observed. Yellow: osmotic stress, in which the cells showed a 16% reduction in viability compared to control.

8.2.6. References

Alamuri, P. and Maier, R. J. 2004. Methionine sulphoxide reductase is an important antioxidant enzyme in the gastric pathogen Helicobacter pylori. *Mol. Microbiol.* **53**, 1397–1406.

Aleman, M., Spier, S. J., Wilson, W. D. and Doherr, M. 1996. Corynebacterium pseudotuberculosis infection in horses: 538 cases (1982-1993). *J. Am. Vet. Med. Assoc.* **209**, 804–809.

Aquino de Sá, M. da C., Gouveia, G. V., Krewer, C. da C., Veschi, J. L. A., de Mattos-Guaraldi, A. L. and da Costa, M. M. 2013. Distribution of PLD and FagA, B, C and D genes in Corynebacterium pseudotuberculosis isolates from sheep and goats with caseus lymphadenitis. *Genet. Mol. Biol.* **36**, 265–268.

Arsenault, J., Girard, C., Dubreuil, P., Daignault, D., Galarneau, J.-R., Boisclair, J., Simard, C. and Bélanger, D. 2003. Prevalence of and carcass condemnation from maedi–visna, paratuberculosis and caseous lymphadenitis in culled sheep from Quebec, Canada. *Prev. Vet. Med.* **59**, 67–81.

Asano, R. L. and Davies, J. 1998. Molecular characterization of the thioredoxin system of Mycobacterium smegmatis. *Res. Microbiol.* **149**, 567–576.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., et al. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75.

Baird GJ, Fontaine MC. 2007.Corynebacterium pseudotuberculosis and its role in ovine caseous lymphadenitis. *J Comp Pathol*, **137**:179–210.

Bandyopadhyay, P. and Steinman, H. M. 2000. Catalase-peroxidases of Legionella pneumophila: cloning of the katA gene and studies of KatA function. *J. Bacteriol.* **182**, 6679–6686.

Bierne, H., Mazmanian, S. K., Trost, M., Pucciarelli, M. G., Liu, G., Dehoux, P., Jänsch, L., Garcia-del Portillo, F., Schneewind, O., Cossart, P., et al. 2002. Inactivation of the srtA gene in Listeria monocytogenes inhibits anchoring of surface proteins and affects virulence. *Mol. Microbiol.* **43**, 869–881.

Billington, S. J., Esmay, P. A., Songer, J. G. and Jost, B. H. 2002. Identification and role in virulence of putative iron acquisition genes from Corynebacterium pseudotuberculosis. *FEMS Microbiol. Lett.* **208**, 41–45.

Braibant, M., Guilloteau, L. and Zygmunt, M. S. 2002. Functional characterization of Brucella melitensis NorMI, an efflux pump belonging to the multidrug and toxic compound extrusion family. *Antimicrob. Agents Chemother.* **46**, 3050–3053.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., et al. 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55**, 611–622.

Calhoun, L. N. and Kwon, Y. M. 2011. Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in Escherichia coli: a review. *J. Appl. Microbiol.* **110**, 375–386.

Carvalho, D. M., Sá, P. H. de, Castro, T. L. P., Carvalho, R. D., Pinto, A., Gil, D. J. P., Bagano, P., Bastos, B., Costa, L. F. M., Meyer, R., et al. 2014. Reference genes for RT-qPCR studies in Corynebacterium pseudotuberculosis identified through analysis of RNA-seq data. *Antonie Van Leeuwenhoek* **106**, 605–614.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M.-A., Barrell, B. G. and Parkhill, J. 2005. ACT: the Artemis Comparison Tool. *Bioinforma. Oxf. Engl.* **21**, 3422–3423.

Casonato S, Provvedi R, Dainese E, Palù G, Manganelli R. 2014. **Mycobacterium tuberculosis require s the ECF sigma factor SigE to arrest phagosome maturation**. *PloS One*, **9**

Chastanet, A., Derre, I., Nair, S. and Msadek, T. 2004. clpB, a novel member of the Listeria monocytogenes CtsR regulon, is involved in virulence but not in general stress tolerance. *J. Bacteriol.* **186**, 1165–1174.

Chevreux, B., Wetter, T. and Suhai, S. *Genome Sequence Assembly Using Trace Signals and Additional Sequence Information*.

Choi, S. H., Baumler, D. J. and Kaspar, C. W. 2000. Contribution of dps to acid stress tolerance and oxidative stress tolerance in Escherichia coli O157:H7. *Appl. Environ. Microbiol.* **66**, 3911–3916.

Costa, Carlota; Capitán, AG; Karachaliou, Niki; Rosell, Rafael.2013. Comprehensive molecular screening: from the RT-PCR to the RNA-seq.*Transl Lung Cancer Res* **2(2)**, 87–91.

Costerton, J. W., Stewart, P. S. and Greenberg, E. P. 1999. Bacterial biofilms: a common cause of persistent infections. *Science* **284**, 1318–1322.

Craig, J. E., Nobbs, A. and High, N. J. 2002. The extracytoplasmic sigma factor, final sigma(E), is require d for intracellular survival of nontypeable Haemophilus influenzae in J774 macrophages. *Infect. Immun.* **70**, 708–715.

Derveaux, S., Vandesompele, J. and Hellemans, J. 2010. How to do successful gene expression analysis using real-time PCR. *Methods* **50**, 227–230.

Doig, P., Boriack-Sjodin, P. A., Dumas, J., Hu, J., Itoh, K., Johnson, K., Kazmirski, S., Kinoshita, T., Kuroda, S., Sato, T.-O., et al. 2014. Rational design of inhibitors of the bacterial cell wall synthetic enzyme GlmU using virtual screening and lead-hopping. *Bioorg. Med. Chem.*

Dorella, F. A., Pacheco, L. G. C., Oliveira, S. C., Miyoshi, A. and Azevedo, V. 2006. Corynebacterium pseudotuberculosis: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. *Vet. Res.* **37**, 201–218.

Douglas, T., Daniel, D. S., Parida, B. K., Jagannath, C. and Dhandayuthapani, S. 2004. Methionine sulfoxide reductase A (MsrA) deficiency affects the survival of Mycobacterium smegmatis within macrophages. *J. Bacteriol.* **186**, 3590–3598.

Everaert, C., Luypaert, M., Maag, J. L. V., Cheng, Q. X., Dinger, M. E., Hellemans, J. and Mestdagh, P. 2017. Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci. Rep.* **7**,.

Farr, S. B. and Kogoma, T. 1991. Oxidative stress responses in Escherichia coli and Salmonella typhimurium. *Microbiol. Rev.* **55**, 561–585.

Ferreira A, O'Byrne CP, Boor KJ. 2001. Role of B in heat, ethanol, acid, and oxidative stress resistance and during carbon starvation in Listeria monocytogenes. *Appl Environ Microbiol*, **67**:4454–4457.

Follmann, M., Ochrombel, I., Krämer, R., Trötschel, C., Poetsch, A., Rückert, C., Hüser, A., Persicke, M., Seiferling, D., Kalinowski, J., et al. 2009. Functional genomics of pH homeostasis in Corynebacterium glutamicum revealed novel links between pH response, oxidative stress, iron homeostasis and methionine synthesis. *BMC Genomics* **10**, 621.

Galardini, M., Biondi, E. G., Bazzicalupo, M. and Mengoni, A. 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. *Source Code Biol. Med.* **6**, 11.

Gal-Mor, O. and Finlay, B. B. 2006. Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cell. Microbiol.* **8**, 1707–1719.

Giorgio, M., Trinei, M., Migliaccio, E. and Pelicci, P. G. 2007. Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nat. Rev. Mol. Cell Biol.* **8**, 722–728.

Gomes, D. L. R., Martins, C. A. S., Faria, L. M. D., Santos, L. S., Santos, C. S., Sabbadini, P. S., Souza, M. C., Alves, G. B., Rosa, A. C. P., Nagao, P. E., et al. 2009. Corynebacterium diphtheriae as an emerging pathogen in nephrostomy catheter-related infection: evaluation of traits associated with bacterial virulence. *J. Med. Microbiol.* **58**, 1419–1427.

Guimarães A de S, Dorneles EMS, Andrade GI, Lage AP, Miyoshi A, Azevedo V, Gouveia AMG, Heinemann MB. 2011. Molecular characterization of Corynebacterium pseudotuberculosis isolates using ERIC-PCR. *Vet Microbiol*, **153**:299–306.

Haas, B. J., Chin, M., Nusbaum, C., Birren, B. W. and Livny, J. 2012. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* **13**, 734.

Hare, R. F. and Hueffer, K. 2014. Francisella novicida pathogenicity island encoded proteins were secreted during infection of macrophage-like cells. *PloS One* **9**.

Ishikawa, M., Okamoto-Kainuma, A., Matsui, K., Takigishi, A., Kaga, T. and Koizumi, Y. 2010. Cloning and characterization of clpB in Acetobacter pasteurianus NBRC 3283. *J. Biosci. Bioeng.* **110**, 69–71.

Jozefczuk, S., Klie, S., Catchpole, G., Szymanski, J., Cuadros-Inostroza, A., Steinhauser, D., Selbig, J. and Willmitzer, L. 2010. Metabolomic and transcriptomic stress response of Escherichia coli. *Mol. Syst. Biol.* **6**, 364.

Kannan, T. R., Musatovova, O., Gowda, P. and Baseman, J. B. 2008. Characterization of a unique ClpB protein of Mycoplasma pneumoniae and its impact on growth. *Infect. Immun.* **76**, 5082–5092.

Kazmierczak MJ, Wiedmann M, Boor KJ. 2005. Alternative sigma factors and their roles in bacterial virulence. *Microbiol Mol Biol Rev*, **69**:527–543.

Kim, J. K., Lee, H. J., Kikuchi, Y., Kitagawa, W., Nikoh, N., Fukatsu, T. and Lee, B. L. 2013. Bacterial Cell Wall Synthesis Gene uppP Is Require d for Burkholderia Colonization of the Stinkbug Gut. *Appl. Environ. Microbiol.* **79**, 4879–4886.

Ladetto, M., Brüggemann, M., Monitillo, L., Ferrero, S., Pepin, F., Drandi, D., Barbero, D., Palumbo, A., Passera, R., Boccadoro, M., et al. 2014. Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia* **28**, 1299–1307.

Lee J-S, Heo Y-J, Lee JK, Cho Y-H. 2005. KatA, the major catalase, is critical for osmoprotection and virulence in Pseudomonas aeruginosa PA14. *Infect Immun*, **73**:4399–4403.

Liu, H., Yang, C.-L., Ge, M.-Y., Ibrahim, M., Li, B., Zhao, W.-J., Chen, G.-Y., Zhu, B. and Xie, G.-L. 2014. Regulatory role of tetR gene in a novel gene cluster of Acidovorax avenae subsp. avenae RS-1 under oxidative stress. *Evol. Genomic Microbiol.* **5**, 547.

Lourdault, K., Cerqueira, G. M., Wunder, E. A. and Picardeau, M. 2011. Inactivation of clpB in the pathogen Leptospira interrogans reduces virulence and resistance to stress conditions. *Infect. Immun.* **79**, 3711–3717.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K. and Smith, I. 2001. The Mycobacterium tuberculosis ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. *Mol. Microbiol.* **41**, 423–437.

McCarthy, R. R., Mooij, M. J., Reen, F. J., Lesouhaitier, O. and O'Gara, F. 2014. A new regulator of pathogenicity (bvlR) is require d for full virulence and tight microcolony formation in Pseudomonas aeruginosa. *Microbiol. Read. Engl.* **160**, 1488–1500.

McKean, S. C., Davies, J. K. and Moore, R. J. 2007a. Probing the heat shock response of Corynebacterium pseudotuberculosis: the major virulence factor, phospholipase D, is downregulated at 43 degrees C. *Res. Microbiol.* **158**, 279–286.

McKean, S. C., Davies, J. K. and Moore, R. J. 2007b. Expression of phospholipase D, the major virulence factor of Corynebacterium pseudotuberculosis, is regulated by multiple environmental factors and plays a role in macrophage death. *Microbiol. Read. Engl.* **153**, 2203–2211.

McNamara, P. J., Bradley, G. A. and Songer, J. G. 1994. Targeted mutagenesis of the phospholipase D gene results in decreased virulence of Corynebacterium pseudotuberculosis. *Mol. Microbiol.* **12**, 921–930.

McNamara, P. J., Cuevas, W. A. and Songer, J. G. 1995. Toxic phospholipases D of Corynebacterium pseudotuberculosis, C. ulcerans and Arcanobacterium haemolyticum: cloning and sequence homology. *Gene* **156**, 113–118.

Meibom, K. L., Dubail, I., Dupuis, M., Barel, M., Lenco, J., Stulik, J., Golovliov, I., Sjöstedt, A. and Charbit, A. 2008. The heat-shock protein ClpB of Francisella tularensis is involved in stress tolerance and is require d for multiplication in target organs of infected mice. *Mol. Microbiol.* **67**, 1384–1401.

Morita, Y., Kodama, K., Shiota, S., Mine, T., Kataoka, A., Mizushima, T. and Tsuchiya, T. 1998. NorM, a putative multidrug efflux protein, of Vibrio parahaemolyticus and its homolog in Escherichia coli. *Antimicrob. Agents Chemother.* **42**, 1778–1782.

Motohashi, K., Watanabe, Y., Yohda, M. and Yoshida, M. 1999. Heat-inactivated proteins are rescued by the DnaK·J-GrpE set and ClpB chaperones. *Proc. Natl. Acad. Sci.* **96**, 7184–7189.

Nair, S. and Finkel, S. E. 2004. Dps protects cells against multiple stresses during stationary phase. *J. Bacteriol.* **186**, 4192–4198.

Neidhardt, F. C. ed. 1996. *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2 edition. Washington, D.C: ASM Press.

Nishimura, T., Teramoto, H., Toyoda, K., Inui, M. and Yukawa, H. 2011. Regulation of the nitrate reductase operon narKGHJI by the cAMP-dependent regulator GlxR in Corynebacterium glutamicum. *Microbiol. Read. Engl.* **157**, 21–28.

Olson, M. E., Ceri, H., Morck, D. W., Buret, A. G. and Read, R. R. 2002. Biofilm bacteria: formation and comparative susceptibility to antibiotics. *Can. J. Vet. Res. Rev. Can. Rech. Vét.* **66**, 86–92.

Osmundson, J., Dewell, S. and Darst, S. A. 2013. RNA-Seq Reveals Differential Gene Expression in Staphylococcus aureus with Single-Nucleotide Resolution. *PLOS ONE* **8**

Pacheco LGC, Castro TLP, Carvalho RD, Moraes PM, Dorella FA, Carvalho NB, Slade SE, Scrivens JH, Feelisch M, Meyer R, Miyoshi A, Oliveira SC, Dowson CG, Azevedo V. 2012. A role for sigma factor σ(E) in Corynebacterium pseudotuberculosis resistance to nitric oxide/peroxide stress. *Front Microbiol*, **3**:126.

Paterson, G. K. and Mitchell, T. J. 2006. The role of Streptococcus pneumoniae sortase A in colonisation and pathogenesis. *Microbes Infect. Inst. Pasteur* **8**, 145–153.

Peel MM, Palmer GG, Stacpoole AM, Kerr TG. 1997. Human lymphadenitis due to Corynebacterium pseudotuberculosis: report of ten cases from Australia and review. *Clin Infect Dis Off Publ Infect Dis Soc Am*, **24**:185–191.

Pinto, A. C., Ramos, R. T. J., Silva, W. M., Rocha, F. S., Barbosa, S., Miyoshi, A., Schneider, M. P. C., Silva, A. and Azevedo, V. 2012. The core stimulon of Corynebacterium pseudotuberculosis strain 1002 identified using ab initio methodologies. *Integr. Biol.* **4**, 789.

Pinto, A. C., Sá, P. H. C. G. de, Ramos, R. T. J., Barbosa, S., Barbosa, H. P. M., Ribeiro, A. C., Silva, W. M., Rocha, F. S., Santana, M. P., Castro, T. L. de P., et al. 2014. Differential transcriptional profile of Corynebacterium pseudotuberculosis in response to abiotic stresses. *BMC Genomics* **15**, 14.

Potter AJ, Kidd SP, Edwards JL, Falsetta ML, Apicella MA, Jennings MP, McEwan AG. 2009.Thioredoxin reductase is essential for protection of Neisseria gonorrhoeae against killing by nitric oxide and for bacterial growth during interaction with cervical epithelial cells. *J Infect Dis*, **199**:227–235.

Radostits OM, Gay CC, Blood DC, Hinchcliff KW.2002. *Clínica veterinária: um tratado de doenças dos bovinos, ovinos, suínos, caprinos e eqüinos* [Veterinary Medicine: A textbook of the diseases of cattle, horses, sheep, pigs and goats]. Guanabara Koogan.

Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., Gallegos, M. T., Brennan, R. and Tobes, R. 2005. The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev. MMBR* **69**, 326–356.

Reen, F. J., Haynes, J. M., Mooij, M. J. and O'Gara, F. 2013. A non-classical LysR-type transcriptional regulator PA2206 is require d for an effective oxidative stress response in Pseudomonas aeruginosa. *PloS One* **8**,

Rodrigue, S., Provvedi, R., Jacques, P.-étienne, Gaudreau, L. and Manganelli, R. 2006. The sigma factors of Mycobacterium tuberculosis. *FEMS Microbiol. Rev.* **30**, 926–941.

Ruiz JC, D'Afonseca V, Silva A, Ali A, Pinto AC, Santos AR, Rocha AAMC, Lopes DO, Dorella FA, Pacheco LGC, Costa MP, Turk MZ, Seyffert N, Moraes PMRO, Soares SC, Almeida SS, Castro TLP, Abreu VAC, Trost E, Baumbach J, Tauch A, Schneider MPC, McCulloch J, Cerdeira LT, Ramos RTJ, Zerlotini A, Dominitini A, Resende DM, Coser EM, Oliveira LM, et al. 2011. Evidence for reductive genome evolution and lateral acquisition of virulence functions in two Corynebacterium pseudotuberculosis strains. *PLoS ONE,* **6**

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A. and Barrell, B. 2000. Artemis: sequence visualization and annotation. *Bioinforma. Oxf. Engl.* **16**, 944–945.

Sansom, F. M., Tang, L., Ralton, J. E., Saunders, E. C., Naderer, T. and McConville, M. J. 2013. Leishmania major methionine sulfoxide reductase A is require d for resistance to oxidative stress and efficient replication in macrophages. *PloS One* **8**.

Sasindran, S. J., Saikolappan, S. and Dhandayuthapani, S. 2007. Methionine sulfoxide reductases and virulence of bacterial pathogens. *Future Microbiol.* **2**, 619–630.

Schumann, W. 2007. Thermosensors in eubacteria: role and evolution. *J. Biosci.* **32**, 549–557.

Sendler, E., Johnson, G. D. and Krawetz, S. A. 2011. Local and global factors affecting RNA sequencing analysis. *Anal. Biochem.* **419**, 317–322.

Shiber A, Ravid T. 2014. Chaperoning proteins for destruction: diverse roles of hsp70 chaperones and their co-chaperones in targeting misfolded proteins to the proteasome. *Biomolecules*, **4**:704–724.

Shin D-H, Choi Y-S, Cho Y-H. 2008.Unusual properties of catalase A (KatA) of Pseudomonas aeruginosa PA14 are associated with its biofilm peroxide resistance. *J Bacteriol*, **190**:2663–2670.

Shinnick, T. M., King, C. H. and Quinn, F. D. 1995. Molecular biology, virulence, and pathogenicity of mycobacteria. *Am. J. Med. Sci.* **309**, 92–98.

Silva, W. M., Seyffert, N., Santos, A. V., Castro, T. L. P., Pacheco, L. G. C., Santos, A. R., Ciprandi, A., Dorella, F. A., Andrade, H. M., Barh, D., et al. 2013a. Identification of 11 new exoproteins in Corynebacterium pseudotuberculosis by comparative analysis of the exoproteome. *Microb. Pathog.* **61**–**62**, 37–42.

Silva, W. M., Seyffert, N., Ciprandi, A., Santos, A. V., Castro, T. L. P., Pacheco, L. G. C., Barh, D., Le Loir, Y., Pimenta, A. M. C., Miyoshi, A., et al. 2013b. Differential Exoproteome analysis of two corynebacterium pseudotuberculosis biovar Ovis strains isolated from goat (1002) and sheep (C231). *Curr. Microbiol.* **67**, 460–465.

Smith, A. W. 2005. Biofilms and antibiotic therapy: is there a role for combating bacterial resistance by the use of novel drug delivery systems? *Adv. Drug Deliv. Rev.* **57**, 1539–1550.

Soares, S. C., Abreu, V. A. C., Ramos, R. T. J., Cerdeira, L., Silva, A., Baumbach, J., Trost, E., Tauch, A., Hirata, R., Mattos-Guaraldi, A. L., et al. 2012. PIPS: pathogenicity island prediction software. *PloS One* **7**

Soares SC, Silva A, Trost E, Blom J, Ramos R, Carneiro A, Ali A, Santos AR, Pinto AC, Diniz C, Barbosa EGV, Dorella FA, Aburjaile F, Rocha FS, Nascimento KKF, Guimarães LC, Almeida S, Hassan SS, Bakhtiar SM, Pereira UP, Abreu VAC, Schneider MPC, Miyoshi A, Tauch A, Azevedo V. 2013. The pan-genome of the animal pathogen Corynebacterium pseudotuberculosis reveals differences in genome plasticity between the Biovar Ovis and Equi strains. *PLoS ONE*, **8**

Squires, C. L., Pedersen, S., Ross, B. M. and Squires, C. 1991. ClpB is the Escherichia coli heat shock protein F84.1. *J. Bacteriol.* **173**, 4254–4262.

Sun, R., Converse, P. J., Ko, C., Tyagi, S., Morrison, N. E. and Bishai, W. R. 2004. Mycobacterium tuberculosis ECF sigma factor sigC is require d for lethality in mice and for the conditional expression of a defined gene set. *Mol. Microbiol.* **52**, 25–38.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J. and Pachter, L. 2010. Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat. Biotechnol.* **28**, 511–515.

Tusher, V. G., Tibshirani, R. and Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 5116–5121.

Vanni, S., Campomanes, P., Marcia, M. and Rothlisberger, U. 2012. Ion binding and internal hydration in the multidrug resistance secondary active transporter NorM investigated by molecular dynamics simulations. *Biochemistry (Mosc.)* **51**, 1281–1287.

Williamson, L. H. 2001. Caseous lymphadenitis in small ruminants. *Vet. Clin. North Am. Food Anim. Pract.* **17**, 359–71.

Zhang, W., Jones, V. C., Scherman, M. S., Mahapatra, S., Crick, D., Bhamidi, S., Xin, Y., McNeil, M. R. and Ma, Y. 2008. Expression, essentiality, and a microtiter plate assay for mycobacterial GlmU, the bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase. *Int. J. Biochem. Cell Biol.* **40**, 2560–2571.

Zhao, H., Sun, Y., Peters, J. M., Gross, C. A., Garner, E. C. and Helmann, J. D. 2016 Depletion of Undecaprenyl Pyrophosphate Phosphatases Disrupts Cell Envelope Biogenesis in Bacillus subtilis. *J. Bacteriol.* **198**, 2925–2935.

Zolkiewski M. 1999. ClpB Cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation: a novel multi-chaperone system from Escherichia coli. *J Biol Chem*, **274**:28083–28086.

8.2.7. Tables

Table 1- Target genes used in the qRT-PCR experiment and the respective primer sequences. The control used was the *dnaG* gene

Gene	Primer Forward '5-3'	Primer Reverse 5'-3'
dnaG (controle)	CCGCTCTATCTTTGTCCCTTATC	TCGTTGACGTTGTCCCATAC
msrA	GGGCAAGAGGTAGTTTATGTGG	CCCAGTACAAACCTCACGATAG
dps	GGCTCACCGACTATAACGATC	TCTTTCTGCTACTTCATCTGCG
tetR2	TCCAATGTCACGGTTGAGG	GCTGTGGAGAGGAAAGACG
glmU	CAACATGACGGTAGGAGAAGG	CCTCCACAAATCCGCCTAAT
norM	CTTCGTAGGAACTGTTTTGGC	ATGATAAGGACCAACGGGATG
clpB	ACCGCAAGTACATCGAGAAAG	TCCTGAATACGCACACCATG
malK	TTACGATTTCCCTGCTAACGAG	GTAATGGCCGCACGAGTAG
cydA	TGTGTCTGGTTCTGGTTGATG	GGTCGGTTGAGGATTTGTCT
sigE	CTTTCACCCAGCTCAAGCC	TCGATCCGATACCAGCGTC

Table 1: Target genes used in the qRT-PCR experiment and the respective primer sequences. The control used was the *dnaG* gene

	Control	pH	2 M
Raw data	23,953.728	14,435.609	12,468.704
Uniquely mapped regions	15,271.603	11,851.744	1,182.943

Table 2 - Number of transcripts obtained under each condition tested. The raw data include readings of all transcripts and uniquely mapped regions of the *C. pseudotuberculosis* 258 genome per condition. Control: no stress; pH: acid stress; 2 M: osmotic stress.

8.2.8. Figures

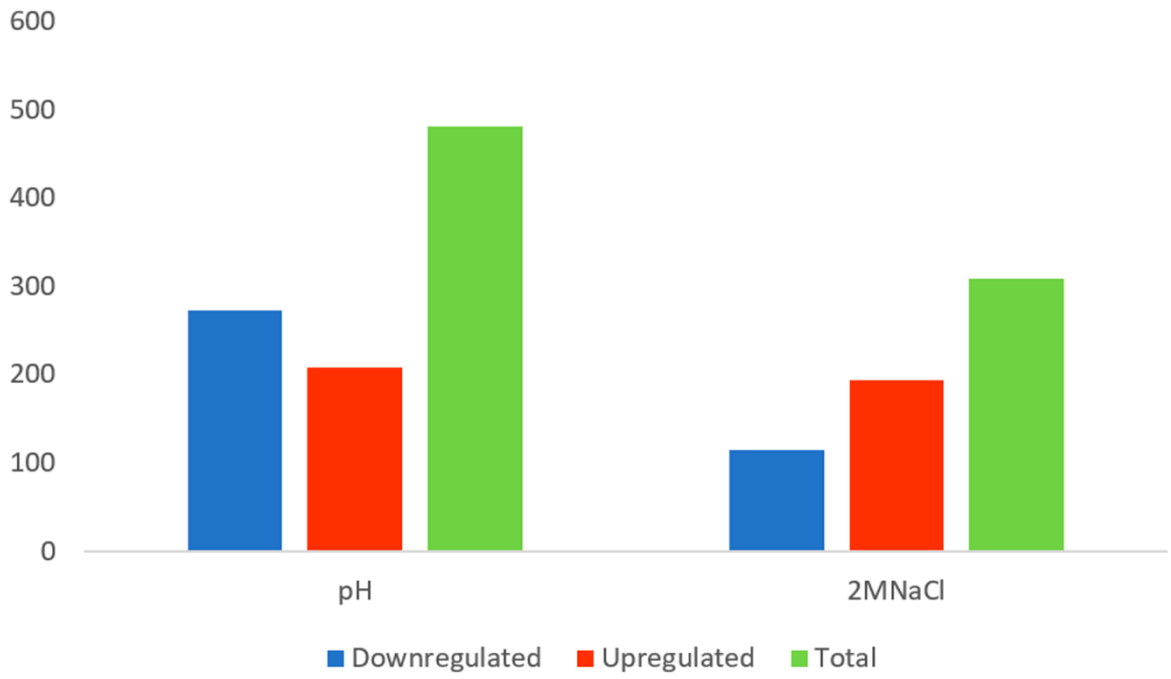


Fig. 1 - Distribution of differentially expressed genes
Under acid stress conditions, 272 genes were considered induced and 208 repressed, totaling 480 genes. Under osmotic stress conditions, 115 genes were induced, and 194 genes were repressed, totaling 309 genes.

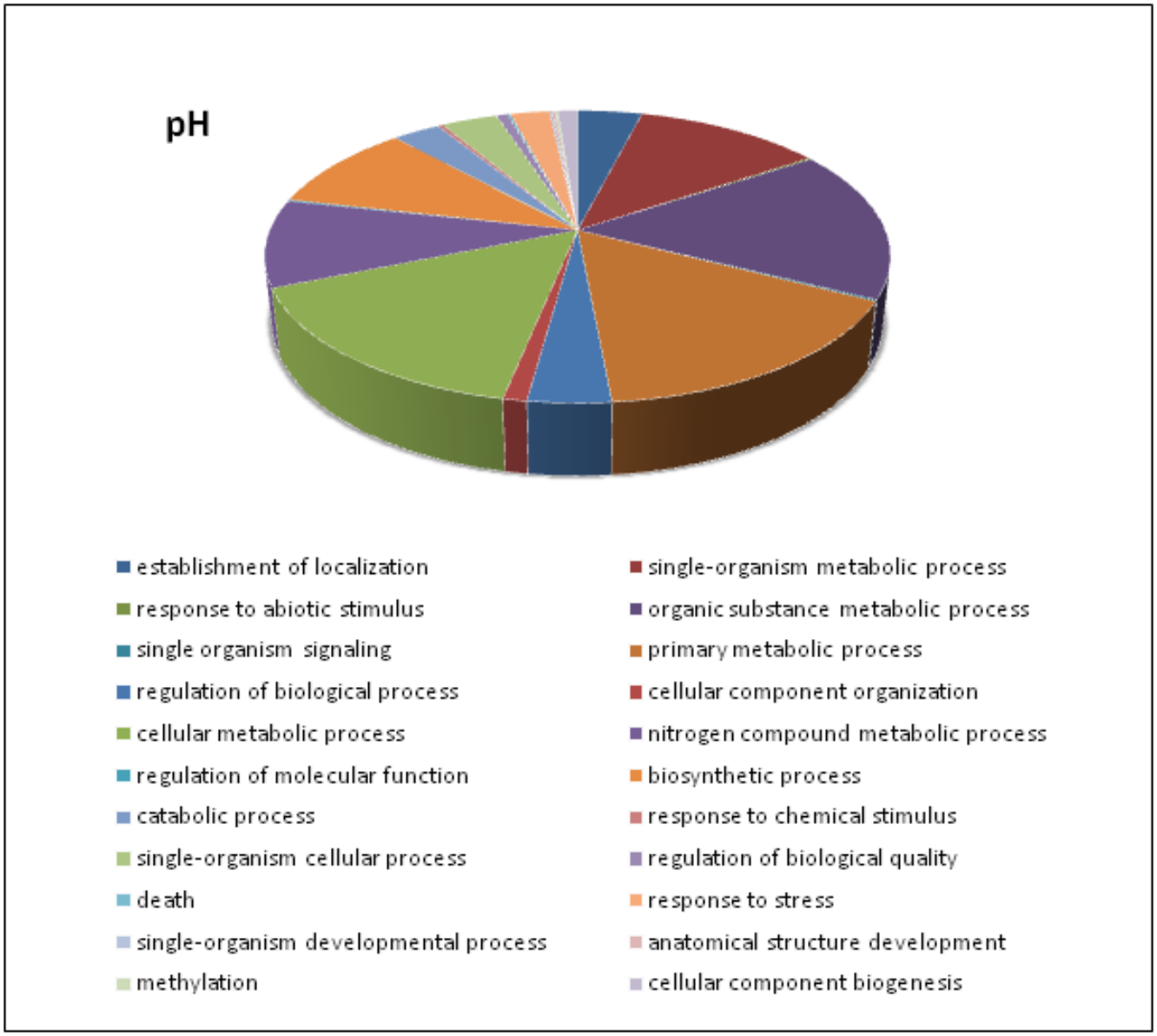


Fig. 2 - Distribution of biological processes involved in acid stress.

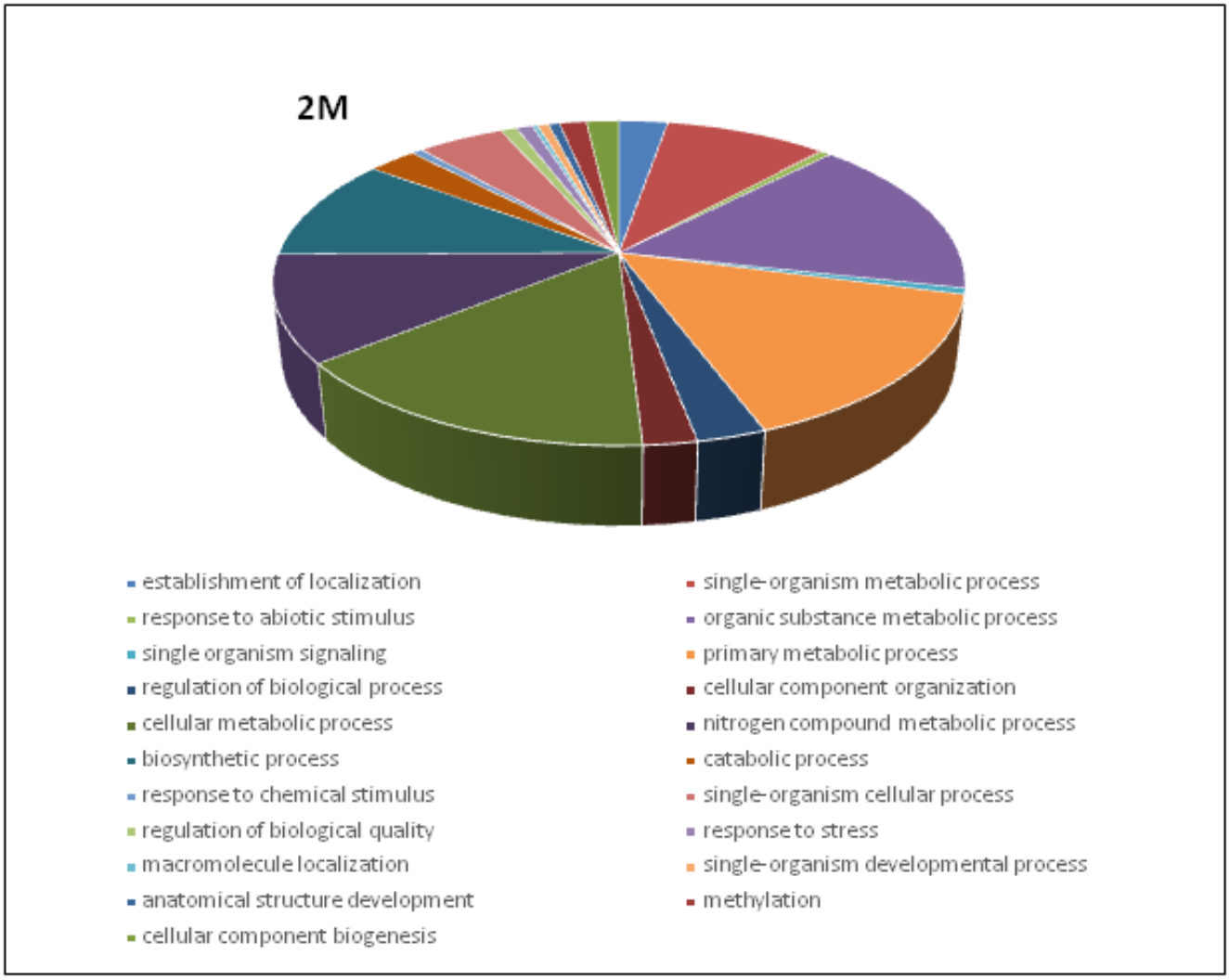


Fig. 3 - Distribution of biological processes involved in osmotic stress.

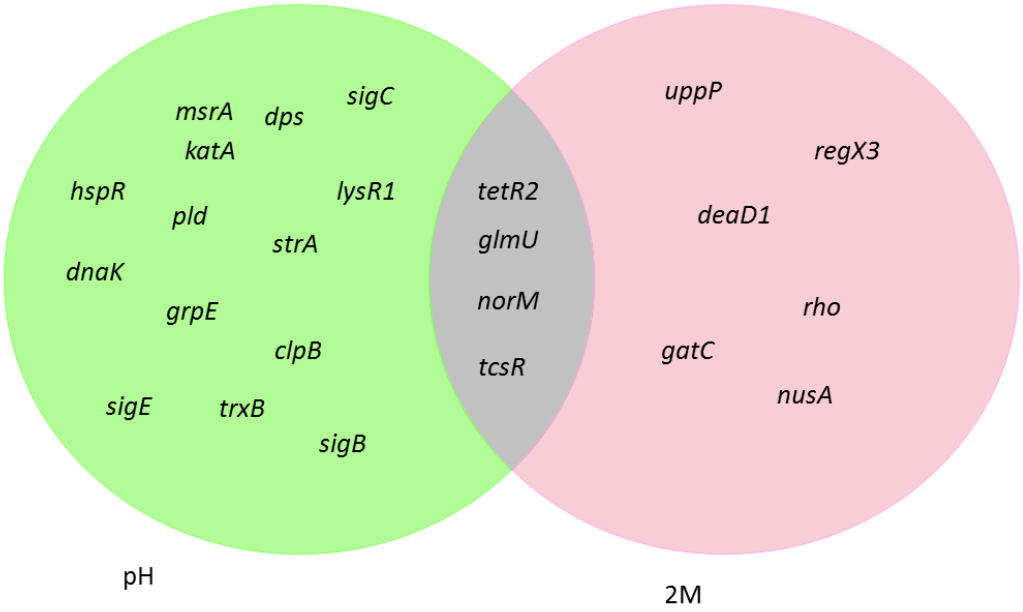


Fig. 4 - Venn diagram of genes, related in the text, composing the biological processes associated with each stress condition. Distribution of genes classified in the category "Biological Processes" and described in the text among the simulated conditions.

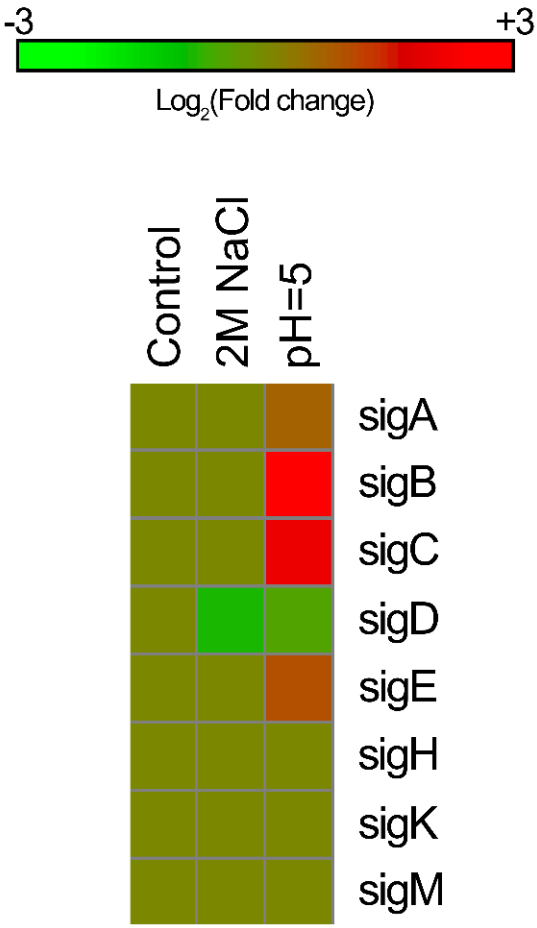


Fig. 5 - Heat map of sigma factors. The color map represents the level of expression of each sigma factor under each stress condition compared to the control condition. The scale shows the values expressed as log 2 (fold change).

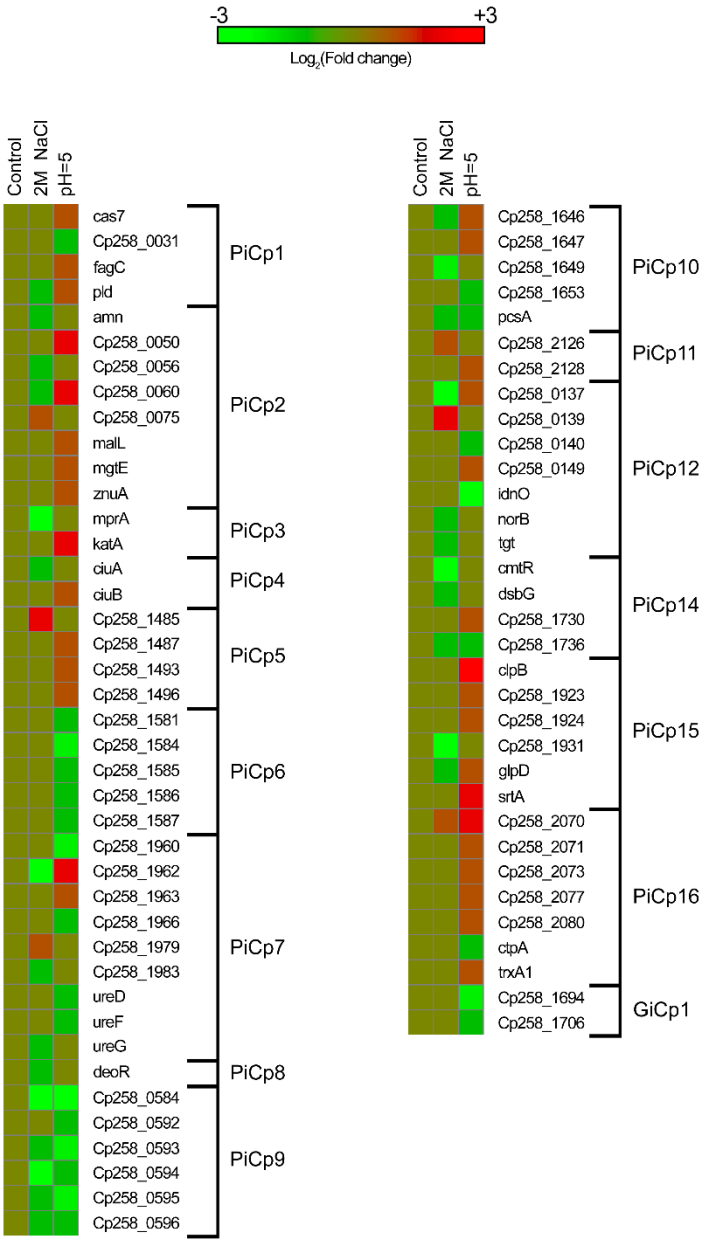


Fig. 6 - Heat map of differentially expressed genes present in pathogenicity islands. The columns show the conditions tested and the rows show the genes associated with each condition. Genes considered induced are shown in red, and genes considered repressed are shown in green. Genes with intermediate expression are shown in dark green. The scale values are expressed as log 2 (fold change).

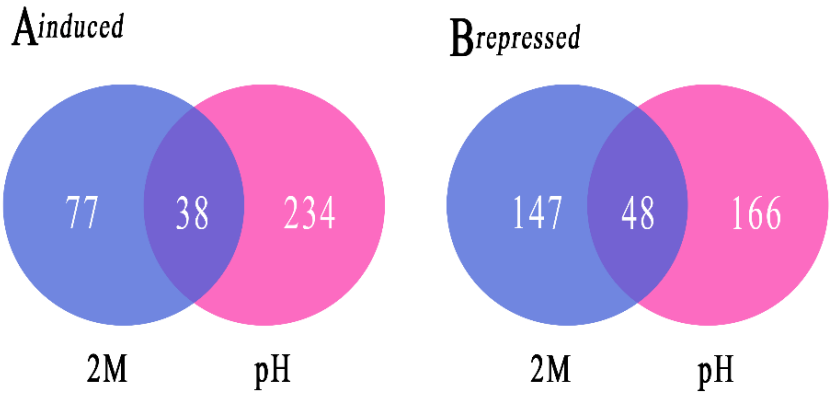


Fig.7 -Venn diagram showing number of shared and unique genes of each simulated condition. Distribution of differentially expressed genes (A, induced; B, repressed) with fold-change ≥ 2 or ≤ 0.5 compared to the control. 2 M: osmotic stress; pH: acid stress.

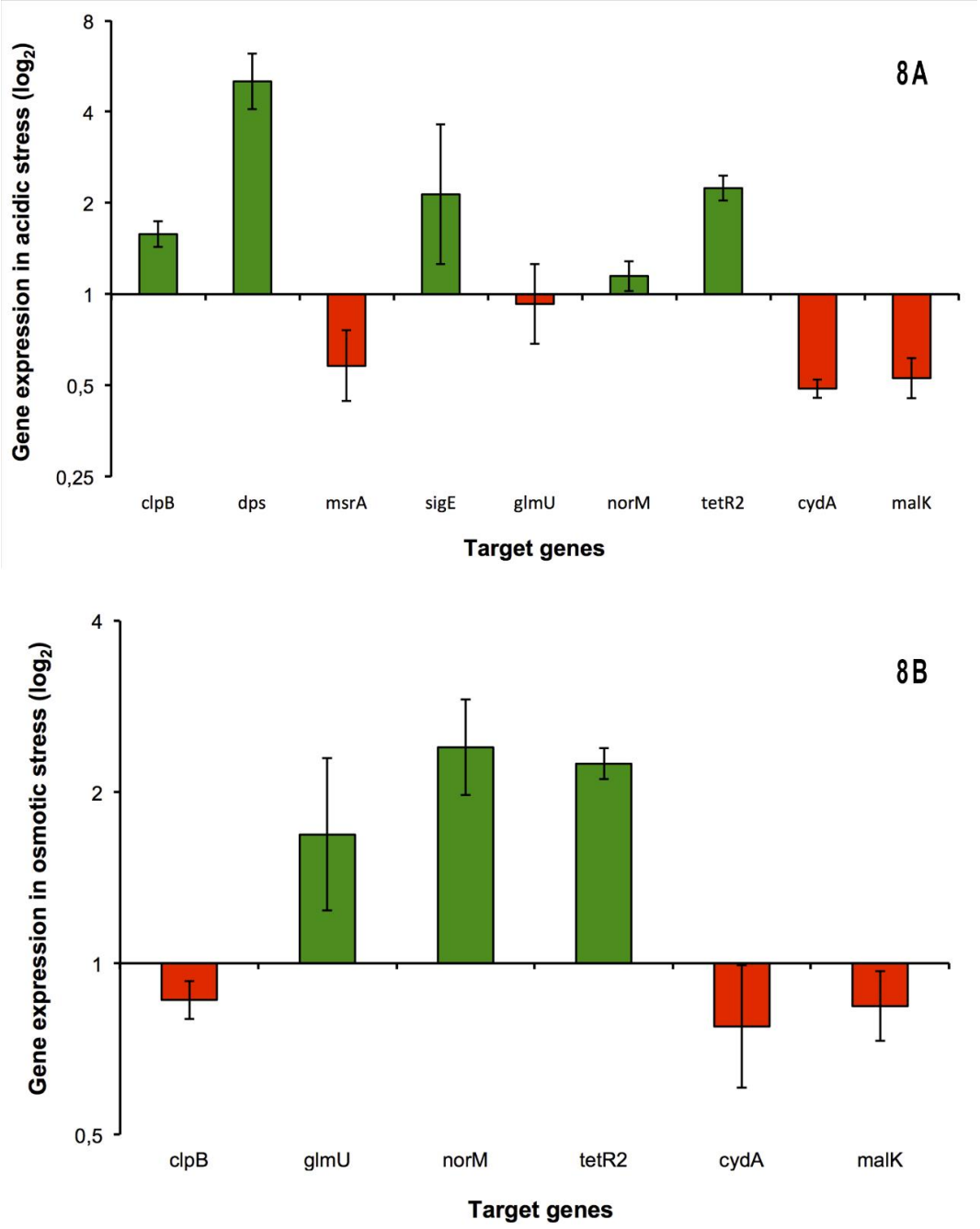
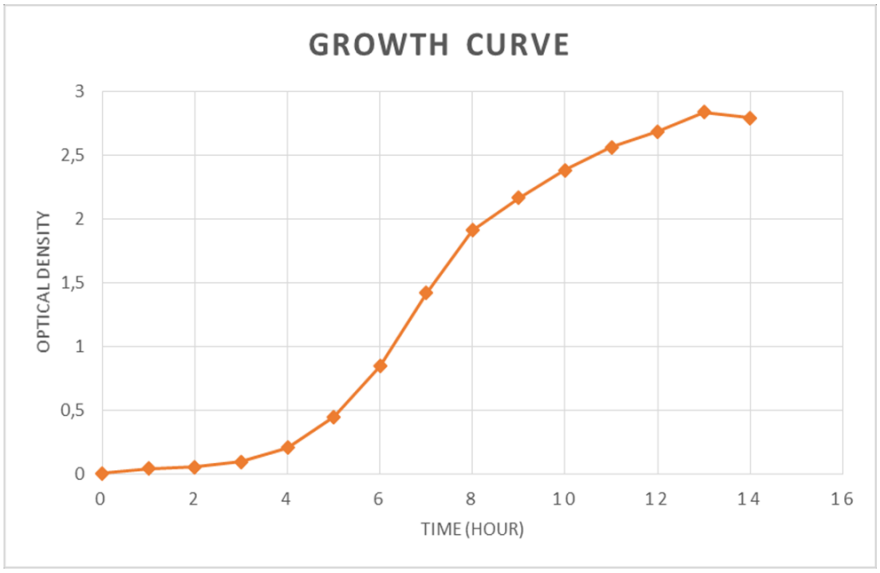


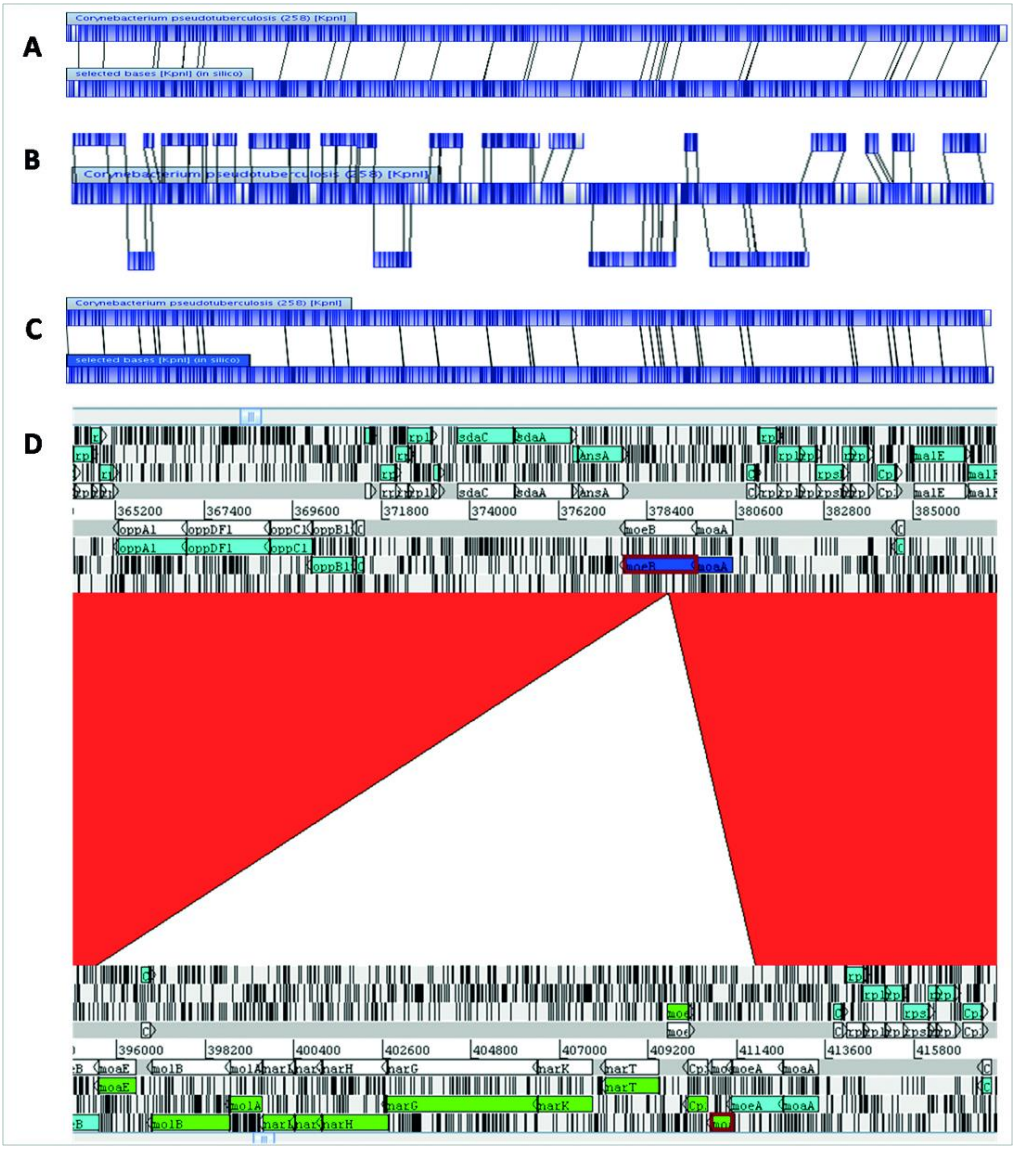
Fig.8 – A. Distribution of the fold change values of the genes obtained in the qRT-PCR experiment in the acid stress. B. Distribution of the fold change values of the genes obtained in the qRT-PCR experiment in the osmotic stress

8.3. ANEXO III: Material suplementar do artigo apresentado no ANEXO II



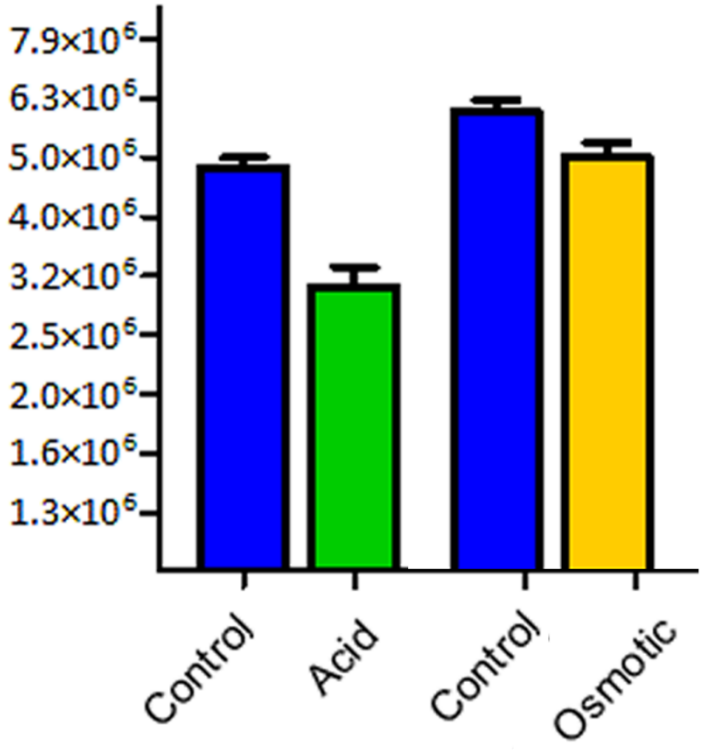
Additional file 1: **Growth curve of *Corynebacterium pseudotuberculosis* strain 258**.

The graph shows the optical density at 600 nm (A600nm) of the culture grown under control conditions for 14 hours. The squares indicate the absorbance values obtained at each time point. Collection of cells for application of stress conditions was performed at OD600nm = 0.2, 2.5 hours after the first measurement.



Additional file 2: *Corynebacterium pseudotuberculosis* 258 genome assembly analysis.

Corynebacterium pseudotuberculosis 258 genome assembly was performed using optical approach whole genome mapping (WGM). Blank regions in the MapSolver software correspond to regions with no alignment; alignments are shaded in blue. (A) Genome sequence of *C. pseudotuberculosis* 258 obtained by sequencing a fragment library with the next-generation genome sequencer SOLiD V3 plus (2,314,404 bp); the orientation of the assembly when compared to the optical map was in the correct order. (B) Scaffolding process of contigs obtained by sequencing simple fragments in the platform Ion PGM™ System for the genome sequence of *C. pseudotuberculosis* 258. (C) Comparison between the Whole-genome mapping with the new genome assembly obtained by Ion PGM™ sequencing. The Fig. shows that the new genome assembly has a length close to the indicated by WGM. In addition, we point out that in comparison with the first published version of the genome, the new assembly of C. *pseudotuberculosis* 258 present an increase of 55,413 bp in the genome length. (D) Synteny graph based on data obtained using ACT between *Corynebacterium pseudotuberculosis* 258 (SOLiD data) (above) and *Corynebacterium pseudotuberculosis* data Ion PGM™ System (below), which shows a 17,000-bp insertion region. Key genes of the *narKGHJI* cluster are mapped to this region.



Additional file 3 - Number of viable cells under each condition tested.

Number of colonies obtained in the colony-forming unit experiment (viability). Blue: control condition. Green: acid stress, in which a 36% reduction compared to control is observed. Yellow: osmotic stress, in which the cells showed a 16% reduction in viability compared to control.