Transcriptome analysis of *Corynebacterium pseudotuberculosis* biovar Equi in two conditions of the environmental stress


**Department of General Biology, Institute of Biological Sciences, Federal University of Minas Gerais, Av. Antônio Carlos, Belo Horizonte 31.270-901, Brazil**

**Laboratory of Genomic and Bioinformatics, Center of Genomics and System Biology, Institute of Biological Science, Federal University of Para, Belém, Pará, Brazil, Rua Augusto Corrêa, Belém 66.075-110, Brazil**

**Federal Rural University of Amazonia, Rodovia PA 140, 2428 Tomé-Açu, PA, Brazil**

**National Institute of Agricultural Technology, Los Reseros y Nicolás Repetto, Hurlingham 1686, Argentina**

**Biotechnology Center, Federal University of Paraíba, João Pessoa, Brazil**

**Department of Computer Sciences, Institute of Exact Sciences, Federal University of Minas Gerais, Av. Antônio Carlos, Belo Horizonte 31.270-901, Brazil**

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

**Keywords:**
Transcriptomics
*Corynebacterium pseudotuberculosis*
RNA-seq
Stress
Biovar Equi
RT-qPCR

**Abstract**

*Corynebacterium pseudotuberculosis* has been widely studied in an effort to understand its biological evolution. Transcriptomics has revealed possible candidates for virulence and pathogenicity factors of strain 1002 (biovar Ovis). Because *C. pseudotuberculosis* is classified into two biovars, Ovis and Equi, it was interesting to assess the transcriptional profile of biovar Equi strain 258, the causative agent of ulcerative lymphangitis. The genome of this strain was re-sequenced; the reassembly was completed using optical mapping technology, and the sequence was subsequently re-annotated. Two growth conditions that occur during the host infection process were simulated for the transcriptome: the osmotic and acid medium. Genes that may be associated with the microorganism's resilience under unfavorable conditions were identified through RNAseq, including genes present in pathogenicity islands. The RT-qPCR was performed to confirm the results in biological triplicate for each condition for some genes. The results extend our knowledge of the factors associated with the spread and persistence of *C. pseudotuberculosis* during the infection process and suggest possible avenues for studies related to the development of vaccines, diagnosis, and therapies that might help minimize damage to agribusinesses.

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 and Fontaine, 2007; Dorella et al., 2006) and a genome with approximately...
50% G + C content with an average size of 2.3 Mb. This bacterium has two biovars, Ovis and Equi, they are classified based on the ability to express the nitrate reductase enzyme; considering that biovar Ovis is negative and biovar Equi is positive for nitrate reduction (Rui et al., 2011). In some bacteria, the presence of genes related to nitrate respiration confers stress resistance inside the host, inferring an association of nitrate respiration with virulence and pathogenicity (Tan et al., 2010). Nonetheless, in C. pseudotuberculosis, the role of nitrate respiration is not known.

C. pseudotuberculosis infects a variety of hosts, including cameldids, cattle, horses, buffaloes, and humans (Alemán et al., 1996; Dorella et al., 2006; Peel et al., 1997). There are many infections caused by this pathogen, among them there is ulcerative lymphangitis caused by the Equi biovar, which mainly affects horses, results in large losses to agribusinesses (Alemán et al., 1996; Peel et al., 1997). In which, C. pseudotuberculosis attacks lymph vessels, producing abscesses scattered throughout the body (Radostits et al., 2002). Thus, it is 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 (H₂O₂). 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, sigma factors are essential for the activation of genes directly or indirectly related to virulence that is, somehow, regulated by the extracytoplasmic function (ECF) (Kazmierczak et al., 2005). Hence, adaptive responses are generated in response to various environmental changes. The mutation in the gene encoding sigma factor E in C. pseudotuberculosis (msrB) attacks lymph vessels, producing abscesses scattered throughout the body (Radostits et al., 2002). Thus, it is 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 (H₂O₂). 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, sigma factors are essential for the activation of genes directly or indirectly related to virulence that is, somehow, regulated by the extracytoplasmic function (ECF) (Kazmierczak et al., 2005). Hence, adaptive responses are generated in response to various environmental changes. The mutation in the gene encoding sigma factor E in C. pseudotuberculosis (msrB) attacks lymph vessels, producing abscesses scattered throughout the body (Radostits et al., 2002). Thus, it is a microorganism of veterinary significance (Arsenault et al., 2003; Guimarães et al., 2011; Williamson, 2001).

The reads were subjected to de novo assembly and to regulation of responses to various stresses (Casonato et al., 2014; Craig et al., 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 et al., 2013a). 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 et al., 2014, 2012). Therefore, the use of this technology has contributed in identifying new possible candidate virulence factors.

Among the most studied genes of C. pseudotuberculosis, there are pld (phospholipase D) and fagB (iron siderophore binding protein) genes, whose functions are essential for the development of lymphangitis. The PLE protein is an exotoxin that aids the organism dissemination from the infection site to the lymph nodes (McKeon et al., 2007a; McNamara et al., 1994, 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 et al., 2002).

The RNAseq tool has increasingly contributed to transcriptome studies in prokaryotes because of its advantages related to other expression analysis technologies, making it an attractive technology for expression profiling studies (Haas et al., 2012; Osmundson et al., 2013).

However, the real-time PCR methodology is still used to validate many of the results obtained, being considered the gold standard for the quantification of differently expressed genes (Derveaux et al., 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 et al., 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 intended to treat, reduce and/or eradicate ulcerative lymphangitis in animals.

2. Materials and methods

2.1. Genomics

2.1.1. 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 KpnI 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.

2.1.2. 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 et al., 1999). 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 first version 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.ich.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.

2.2. Transcriptomics

2.2.1. Bacterial cultures

Corynebacterium pseudotuberculosis biovar Equi strain 258, isolated from a horse in Belgium, was cultivated in plates with Brain Heart Infusion Broth (BHI) consisting of 200 g of 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, 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 37°C in a shaker at...
160 rpm. Following this period, a 1/10 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 h (see Additional file 1).

2.2.2. Application of stress conditions

The experiment was carried out at the beginning of the exponential phase, so 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 (A600 = 0.2), it was divided into two Falcon tubes (20 mL each), which were centrifuged for 3 min at 8000 rpm at 37 °C. 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 from 10−1 to 10−6. The dilutions 10−4, 10−5 and 10−6 were plated on BHI agar, and the plates were incubated at 37 °C for 48 h followed by a determination (in duplicate) of cell viability and colony counting. The remnants of each sample were centrifuged at room temperature for 3 min at 8000 rpm, and the pellet was resuspended in 2 mL RNA later® (Ambion, USA) according to the manufacturer’s instructions.

2.2.3. 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 microtubes containing the samples, and the tubes were homogenized using a Precellys 24 homogenizer at 6500 rpm for two 15-second cycles with a 30-second interval between cycles. The samples were then centrifuged for 1 min, and the supernatant was transferred to a new 2-mL tube and incubated in a dry bath at 60 °C for 15 min. DNase was added to avoid any contamination with genomic DNA. The total RNA extracted was eluted in 100 μL of RNA-free milli-Q water and quantified using a Qubit® 2.0 fluorometer (Invitrogen, USA).

2.2.4. rRNA depletion for mRNA enrichment

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

2.2.5. 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, USA).

The RNA was fragmented through the RNase III enzyme, adding end adapters, and the sample was amplified via reverse transcriptase, following the protocol provided by the SOLiD™ Total RNA Seq kit (Life Technologies™, CA). After, fragments of the cDNAs amplified by PCR in the 150 to 250 base pair range were selected via electrophoresis on 6% polyacrylamide gel. 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 caught and deposited on slides to perform the sequencing on the SOLID™ V3 platform, according to the manufacturer’s recommendations.

2.2.6. Transcriptomic analysis

Two files were obtained following sequencing: one in .csfasta format and another in QV qual format; these files contained the reads generated in color space and the quality of the bases in 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 > 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 v2.0.2 (Sasindran et al., 2007). 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 average 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.blast2go.com), which used a gene ontology database.

2.2.7. RT-qPCR assay

The total bacterial RNA was reversed transcribed into cDNA using the High-Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). cDNA synthesis were carried out in the ATC 401 (Nyx Technik) thermocycler and samples were stored in the −80°C 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

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Target genes used in the qRT-PCR experiment and the respective primer sequences. The control used was the dnaG gene.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene</strong></td>
<td><strong>Primer Forward 5′-3′</strong></td>
</tr>
<tr>
<td>dnaG (control)</td>
<td>CCGCTCTATCTTCTGTTCCTTAGTC</td>
</tr>
<tr>
<td>murA</td>
<td>GCCGAAAGGATGGTGTGTTAAGGG</td>
</tr>
<tr>
<td>rpoB</td>
<td>GGCTCAAGCAGACTAAGAGGAT</td>
</tr>
<tr>
<td>glmU</td>
<td>CAACTGACGGGTAGGGAAGG</td>
</tr>
<tr>
<td>norM</td>
<td>CTGGTAGAAGAAGTTTGGG</td>
</tr>
<tr>
<td>cdpB</td>
<td>ACCCGAGGATGAGAGAAGG</td>
</tr>
<tr>
<td>madK</td>
<td>TTACGATTCCCCGGGTTAGAG</td>
</tr>
<tr>
<td>cydA</td>
<td>TGTGTCCTGGTCTGCTGTTAG</td>
</tr>
<tr>
<td>sigE</td>
<td>CTTTCACCCATGCTCAAAGC</td>
</tr>
</tbody>
</table>
amplification efficiency between 90 and 100% (Table 1). Analysis were performed in triplicates for each 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).

3. Results and discussion

3.1. Genomics

3.1.1. C. pseudotuberculosis 258 genome refinement

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

The comparison analysis of the results with the first published version of the genome (accession number CP003540.1) showed that the whole-genome map presents 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 > 20, indicating no need for quality filter application. De novo assembly performed using the software Mira (Chevreux et al., 1999) showed 41 contigs, 19.59× average coverage, and N50 value of 93,650. The contigs scaffolding 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 2369.817 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 narKGHIJ, which is responsible for nitrate reduction, were identified (Nishimura et al., 2011). This region was missing in the first version of the sequence.

3.2. Transcriptomics

3.2.1. 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 (A600nm = 0.2). Cell viability was assessed through the colony-forming unit assay. The results showed a 36% decrease in replication under acidic stress, a 16% decrease under osmotic stress conditions (Jozefczuk et al., 2010). 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 and Kogoma, 1991). Thus, bacterial molecules undergo redox reactions with ions and molecules present in the environment to survive, consequently achieving internal homeostasis (Follmann et al., 2009). When cells face this type of adverse conditions, proteins undergo changes that impair their function and require protection. We highlight the following genes based on this profile: mraA, dps in the acidic 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).

3.2.2. Biological processes

Gene ontology (GO) analyses showed that 22 biological processes were involved in the response to acidic stress, and 21 processes were involved in the osmotic stress (Gene Ontology level 3; Figs. 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: Supplementary Table S1).

3.3. 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 and Kogoma, 1991). Thus, bacterial molecules undergo redox reactions with ions and molecules present in the environment to survive, consequently achieving internal homeostasis (Follmann et al., 2009). When cells face this type of adverse conditions, proteins undergo changes that impair their function and require protection. We highlight the following genes based on this profile: mraA, dps in the acidic 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 genes were 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 methionine sulfoxide reductase repair protein is encoded by the genes msrA and msrB. These genes share no homology and display the same enzymatic activity, although operating on different substrates (Sasindran et al., 2007). Only msrA, which showed an 8.7-fold change under acidic stress (Fig. 4) was considered induced in strain 258 (Equi), whereas msrB was induced only under acidic conditions in strain 1002 (Ovis) (Pinto et al., 2014). Previous studies have shown that the role of MsrA is more relevant to virulence than MsrB (Sasindran et al., 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$_2$O$_2$) 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, presented in studies conducted on animal models (Alamuri and Maier, 2004). These results provide strong evidence for the importance of this gene in contributing to the bacteria resilience 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 acidic 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 and Kwon, 2011). Its function contributes to bacterial protection against acidic, oxidative and heat stress, UV exposure, gamma radiation and iron and copper toxicity during the stationary phase (Calhoun and Kwon, 2011; Nair and 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 acidic environment (pH = 1.8) (Choi et al., 2000). The dps gene was also induced under acidic stress in C. pseudotuberculosis 1002 (Pinto et al., 2014), indicating that it is an important candidate for studies related to fighting caseous lymphadenitis and ulcerative lymphangitis.

### 3.4. Induced genes involved in the regulation of biological processes

The regulation of biological processes forms the basis of living beings’ ability to maintain balance 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 acidic 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 lsrR1 and sigB_2E (Fig. 4) genes were present only in the acidic condition and we highlight due to the fold change, but only lsrR1 gene was related in this process. The sigma factor genes were listed in the item about sigma factors. The tetR2 gene, that showed approximately a nine-fold change under osmotic stress conditions and an 18 fold change under acidic 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). A study about 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.

**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.
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 functional 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. Hence, 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 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 for 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 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.

3.5. Genes involved in responses to chemical stimuli

Responses to chemical stimuli refer 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 outstanding because of their expression values, which were higher under the conditions tested (see Additional file 4: Supplementary Table S1).

The trxB gene, showed a 7.2-fold change under acid stress condition and encodes a thioredoxin reductase that, together with thioredoxin (trxA), forms part of the thioredoxin system (Trx) (Asano and 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 et al., 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 required 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 et al., 2002; Morita et al., 1998). The ABC transporters promote the extrusion of toxic molecules and drugs via a proton/electron electrochemical gradient through the membrane, contributing to intracellular homeostasis (Vanni et al., 2012). In C. pseudotuberculosis 1002 (Ovis), norM may also be expressed under osmotic stress conditions with a significant fold change (3.2) (Pinto et al., 2014), favoring its potential use in controlling 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 bacteria, such as E. coli, Burkholderia sp., Bacillus subtilis, among others, it is related to the biosynthesis of different cell wall components and membrane proteins, 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). This gene, 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. Therefore, becoming an important antibacterial target.

3.6. Expression of sigma factor-encoding genes

The genome of C. pseudotuberculosis strain 258, as identified by (Pinto 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, that 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), 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 its 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 et al., 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 required 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 it is, apparently, involved in the bacterial resistance to oxidative agents. The same study also showed that the sigma E factor may affect the expression of the alternative sigma factor, SigH (Manganelli et al., 2001). These genes and their regulons may form part of a gene catalog to be tested in studies involving the development of therapies, diagnosis 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 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 it is one of only two sigma factors with extracytoplasmic function that have been found in M. leprae (Rodríguez 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 it was induced at the beginning of the exponential phase of growth in acid medium.

3.7. 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, expressed when the bacterium comes into contact with the host, and they are located in regions called pathogenicity islands (Gal-Mor and Finlay, 2006; Hare and Hoeff, 2014).

A total of 16 pathogenicity islands were predicted in *C. pseudotuberculosis*, totaling 265 genes (Soares et al., 2013, 2012). 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 H2O2 stress and in virulence (Bandyopadhyay and Steinman, 2000; Lee et al., 2005). The *katA* gene is expressed at all stages of bacterial growth but especially in the stationary phase (Shin et al., 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 et al., 2008). Biofilm formation is related to chronic infection and to resistance to antibiotics (Costerton et al., 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 et al., 1999; Smith, 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 et al., 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 et al., 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 (43 °C), 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 et al., 2007a, 2007b). These studies corroborate our results in C. pseudotuberculosis 258.

The srtA gene was induced under acid stress conditions (7.0-fold change) (Fig. 4). SrtA encodes a group of enzymes called sortases that is found in all gram-positive bacteria and is involved in the ability to cause various diseases (Paterson and 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 srtA 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 and 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), 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 and Ravid, 2014). ClpB is required for cellular tolerance of various stresses, including heat, osmotic, ethanolog, 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.

3.8. 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.

Relating to the stimulons analyzed, the number of hypothetical protein coding sequences was high and exhibited significant fold change, and many belonging to 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 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. Hence, they 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).

3.9. 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 et al., 2013). 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 \textit{C. pseudotuberculosis}, the \textit{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 \textbf{Table 1}. In the acid stress condition, with the exception of the \textit{msrA} that was repressed and the \textit{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™ platform. In the osmotic stress condition, all the genes were confirmed and presented the same expression profile as the one detected by the SOLiD™ 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 bacteria.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gene.2018.08.028.

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.

References


