Spring 5-5-2018

Francisella tularensis Insertion Sequence Elements Contribute to Differential Gene Expression

Amanda M. Bartling
University of Nebraska Medical Center

FRANCISELLA TULARENSIS INSERTION SEQUENCE ELEMENTS CONTRIBUTE TO DIFFERENTIAL GENE EXPRESSION.

by

Amanda Marie Bartling

A THESIS

Presented to the Faculty of

the University of Nebraska Graduate College

in Partial Fulfillment of Requirements

for the Degree of Master of Science

Pathology and Microbiology Graduate Program

Under the Supervision of Professor Marilynn A. Larson

University of Nebraska Medical Center

Omaha, Nebraska

December 2017

Advisory Committee:

Marilynn A. Larson, PhD Steven H. Hinrichs, MD

Paul D. Fey, PhD

FRANCISELLA TULARENSIS INSERTION SEQUENCE ELEMENTS CONTRIBUTE TO DIFFERENTIAL GENE EXPRESSION.

Amanda M. Bartling, M.S.

University of Nebraska Medical Center, 2018

Advisor: Marilynn A. Larson, PhD

Francisella tularensis is the causative agent of the disease tularemia and a select agent. The genome of this pathogen contains many insertion sequence (IS) elements with IS*Ftu1* being the most abundant. IS*Ftu1* expression and the contribution of this IS element to differential gene expression in the *F. tularensis* subpopulations was evaluated. Full-length and truncated IS*Ftu1* sense and antisense transcriptional expression was detected. The prototype A.I strain (Schu S4) had considerably higher expression levels of IS*Ftu1* and the adjacent genes than the wild type A.I strain (NE061598). The A.II strains (WY96-3418 and WY-00W4114) had similar expression levels for IS*Ftu1* and the adjacent genes. In the highly virulent A.I strains, a bicistronic transcript encoding the universal stress protein (Usp) and a downstream IS*Ftu1* was highly expressed during early and mid log growth phase when provisions were plentiful, and was induced 2-fold by nitric oxide and a polyamine. During late log growth and stationary phase, only monocistronic transcripts for *usp* and IS*Ftu1* were being moderately expressed. Intrinsic transcription termination sequences were not apparent between *usp* and IS*Ftu1*. Secondary RNA structure models indicated that the bicistronic transcript will form more readily than the monocistronic *usp* transcript. The co-expression of IS*Ftu1* and the adjacent gene(s) may provide a fitness advantage by regulating expression and/or transcript stability of the co-transcribed gene(s). Additional study will provide a better understanding of the contribution of IS elements to the differential gene expression and virulence observed for the various *F. tularensis* clades.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

MITE	miniature inverted-repeat transposable elements
REP	repetitive extragenic palindromic
IS	insertion sequence
bp	base pairs
kb	kilobases
ORF	open reading frame
BIME	bacterial interspersed mosaic element
LTR	long-terminal repeat
IHF	integration host factor
FPI	*Francisella* pathogenicity island
T6SS	type 6 secretion system
LVS	Live Vaccine Strain
LPS	lipopolysaccharide
TLR	toll-like receptor
CR3	complement receptor 3
FCP	*Francisella* containing phagosome
INF	interferon
iNOS	nitric oxide synthase
SNAP	S-nitroso-acetyl-penicillamine
DISA	differential insertion sequence amplification
PFGE	pulse field gel electrophoresis
CDC	Centers for Disease Control and Prevention
HHS	Department of Health and Human Services

BSL-3	biosafety level 3
BHI	brain heart infusion
cV	constant volts
rRNA	ribosomal RNA
RT-PCR	reverse transcription-polymerase chain reaction
RACE	rapid amplification of cDNA ends
NIH	National Institute of Health
DDBJ	DNA DataBank of Japan
ENA	European Nucleotide Archive
BLAST	Basic Local Alignment Search Tool
NCBI	National Center for Biotechnology Information
NLM	National Library of Medicine
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
IGV	Integrative Genomics Viewer
Mb	million bases
UTR	untranslated region
Usp	universal stress protein
ANI	average nucleotide identity
HtpG	high-temperature protein G

INTRODUCTION

Mobile Genetic Elements in Bacteria

Bacteria face a constant struggle to survive in their environments. They have to be stable enough to pass on their genetic information, but variable enough to adapt to their changing surroundings, including the ability to survive different stresses and evade host immune surveillance. In order to achieve this balance, bacteria utilize several mechanisms to make their genome more plastic. Some of these mechanisms entail the acquisition of mobile genetic elements via horizontal or vertical gene transfer, genomic rearrangements, and/or acquisition of mutations. Mobile genetic elements can be involved with gene transfer, genomic rearrangements, and mutations, as well as the regulation of bacterial genes such as virulence factors. Since mobile genetic elements can influence the phenotype of an organism, including pathogenicity, the study of mobile genetic elements or mobilomics has been developed and is an important area of investigation (Frost et al., 2005). Therefore, this thesis will focus on mobile genetic elements and their contribution to genomic flexibility and effect on gene expression, including the survival of pathogens in diverse ecological niches.

Mobile genetic elements are transposable DNA sequences that are able to excise from the genome and re-insert into a genome at a different location without using homologous recombination. Mobile elements may change location, or they may duplicate and change their copy number. Most of these elements have inverted repeats that allow for recognition and processing. Mobile genetic elements can affect gene expression by disrupting, activating, and /or altering the regulation of the effected gene(s) (Darmon and Leach, 2014). Accordingly, mobile elements can play a role in the regulation of virulence factors, as well as genome reduction (Dobrindt and Hacker, 2001). However, all mobile

genetic elements have some kind of regulation to prevent genomic instability and excessive mutagenesis.

Mobile genetic elements include plasmids, miniature inverted-repeat transposable elements (MITEs), repetitive extragenic palindromic (REP) sequences, genomic islands, transposable bacteriophages, group I introns, group II introns, transposons, and insertion sequence (IS) elements (Darmon and Leach, 2014). Each of these elements will be discussed, with particular focus on bacterial IS elements, along with their role in influencing gene expression and pathogenicity. For a schematic organization of different transposable elements, see Figure 1 in the review by Darmon and Leach (2014). Table 1 in this review provides examples of various mobile genetic elements that were inserted into a prokaryotic genome, as well as the resulting alterations in gene organization and expression.

Plasmids contain genes that encode for their own transfer and can contain other DNA, such as antibiotic resistance genes, that are organized into a self-replicating replicon. Most plasmids are circular double-stranded DNA; however, a few linear double-stranded DNA plasmids have also been discovered (Hinnebusch and Tilly, 1993). Plasmids replicate at each cell division through a process called partitioning. Plasmids are incapable of co-existing with another type of plasmid that has the same replication mechanism in a cell. Plasmids can be transferred via conjugation in which a mating-pair is formed followed by signaling that the transfer of the DNA plasmid can begin and finally the actual transfer of DNA (Frost et al., 2005). Plasmids can spread antibiotic resistance genes and virulence factors. For example, in the case of *Bacillus anthracis,* this select agent pathogen has acquired one plasmid that encodes a capsule (pXO2) and one plasmid that encodes a toxin (pXO1) (Frost et al., 2005; Kolstø et al., 2009).

MITEs are small DNA sequences (100-500 base pairs in length) that are usually AT rich with a TA dinucleotide motif and target site duplications at each end. MITEs do not encode for their own transposase, but mobilization in bacteria has been shown (Bardaji et al., 2011). MITE insertion can add functional open reading frames (ORFs) to the genome, inactivate a gene, or modify the transcription of nearby genes through the addition of a promoter or regulatory binding site (Darmon and Leach, 2014). MITEs are also capable of recombining which can cause genomic rearrangements or deletions of the genome.

REP sequences are often found in pairs or clusters and are usually located in transcribed regions of the genome. They are typically between 21 and 65 bases (Magnusson et al., 2007). REPs are highly conserved and may form stable stem-loop structures (Newbury et al., 1987). A pair of inverse REP sequences separated by a short linker with conserved sequence motifs is called a bacterial interspersed mosaic element (BIME). REP and BIME sequences can play a role in the regulation of gene expression by being transcribed into a stable RNA structure that binds to a transcript, preventing 3'-5' degradation, affecting function, and/or inhibiting translation of the respective protein. There is also evidence that REP sequences may be involved with transcription termination, protein interactions, and some have been shown to stimulate immune function in mammalian cells (Newbury et al., 1987).

Genomic islands are large discrete sequences of DNA (10-200 kb in length), which differ among closely related strains (Darmon and Leach, 2014). Some genomic islands are mobile, while others are not. There are different types of islands depending on the effect of the accessory genes that are encoded, including pathogenicity islands, metabolic islands, ecological islands, or defense islands (Darmon and Leach, 2014). Similar islands may perform different functions in different bacteria, depending on internal and external

signals. These islands are flanked by direct repeats and generally contain a different G+C content than the rest of the genome (Hacker and Kaper, 2000). There are usually mobile elements encoded in the genomic islands, which may be functional or nonfunctional. If the mobile element(s) are functional, genomic islands can be excised and reinserted into a different location in the genome, or they can undergo horizontal gene transfer to another bacterial cell that is not from parent to progeny. In addition, the genes encoded in a genomic island can be self-regulated or other factors within the bacterial genome can contribute to their regulation, in response to internal and/or external cues (Darmon and Leach, 2014).

Bacteriophages are viruses that are able to transpose their DNA into a bacterial genome. When the bacteriophage DNA is inserted into the bacterial chromosome or plasmid, the surrounding sequence is usually duplicated during this process. Bacteriophages, like Mu phages, are integrated randomly into the host genome (Darmon and Leach, 2014). Predictably, this can lead to inactivation of genes that are disrupted by these insertions. This can cause further disruption if those genes are part of an operon. It is also possible that bacteriophages can induce other bacteriophages to become mobile or cause recombination events to occur. Bacteriophages have been shown to uptake host DNA with subsequent injection into another host, but this occurs at a low frequency (Frost et al., 2005).

Introns are self-splicing RNA molecules that are widespread in prokaryotes, but are not very abundant. Introns can insert into a homologous DNA site, called homing, or they can insert into a novel, ectopic DNA site, called transposition (Tourasse and Kolstø, 2008). They are splicing ribozymes as well as mobile elements that are divided into two groups, which are referred to as group I introns and group II introns. Group I introns are transcribed into self-splicing RNAs that have 10 helices capped by loops (Darmon and

Leach, 2014). Group I introns are found in tRNA or rRNA genes and in bacteriophages. They encode an endonuclease that enables their integration into a DNA site. Once integrated into a bacterial genome, some group I introns lose their mobility gene and are unable to splice out.

Although group II introns are self-splicing RNAs, they are more similar to retroelements than group I introns. They contain six double-helical domains and encode for a protein with endonuclease, reverse transcriptase and RNA maturase activities (Darmon and Leach, 2014). Group II introns catalyze their own splicing with help of their encoded protein that stabilizes the RNA structure. These introns then invade DNA using their inherent ribozyme activity, followed by reverse transcription back into the DNA (Lambowitz and Zimmerly, 2011). Many bacterial group II introns have deleted regions or have inserted foreign DNA, and they often integrate into other mobile elements, especially plasmids, which maintains their proliferation. Group II introns can also alter the sequence of the host chromosome by causing deletions, inversions, or genomic rearrangements (Darmon and Leach, 2014). Interestingly, group II introns are believed to be ancestors of nuclear spliceosomal introns in eukaryotes (Tourasse and Kolstø, 2008).

Transposons generally have long inverted repeats at each end and contain genes that provide an advantage to the bacteria. Transposons can also have strong site preferences for insertion into a genome and can induce genomic rearrangement, disrupt genes or change the regulation of neighboring genes (Wang et al., 2006). The upstream open reading frame of the transposase generally encodes the DNA recognition domain, which recognizes specific sequences for DNA binding, and the downstream open reading frame encodes the catalytic site, which catalyzes the transposition reaction (Mahillon and Chandler, 1998). In general, there are two main types of transposons, retrotransposons (Class I) and DNA transposons (Class II) (Hadjiargyrou and Delihas, 2013).

Retrotransposons are further divided into long-terminal repeat (LTR) retrotransposons and non-LTR retrotransposons. Retrotransposons use a copy-and-paste mechanism to move through the genome. They create a RNA intermediate that is then reverse transcribed back into the genome (Beauregard et al., 2008). DNA transposons usually move via a cut-and-paste mechanism whereby they are excised from one position in the genome and integrate into a new position (Muñoz-López and García-Pérez, 2010). A compound transposon is flanked by IS elements and encode a functional transposase. Conjunctive transposons can relocate intracellularly or intercellularly through conjugation (Darmon and Leach, 2014).

IS elements are small DNA regions (0.7-2.5 kb in length) that contain one or two ORFs encoding a transposase. This transposase is what allows the insertion sequence to be mobilized. The terminal regions contain short inverted repeats between 10 and 40 bp (Mahillon and Chandler, 1998). Insertion of an IS element will change the bacterial genome with the addition of the transposase gene and creates direct repeats of the target DNA flanking the IS at the site of insertion. Depending on the location of the insertion, IS elements can induce phase variation, modify gene expression, or disrupt genes.

Insertion Sequence Elements

IS elements recognize specific sites for recombination, and many IS elements contain a similar motif in the transposase catalytic domain. This motif is referred to as the DDE motif (Mahillon and Chandler, 1998). This motif can also be found in retroviruses, eukaryotic transposons, and retrotransposons. Mutations in these sites have been shown to inactivate transposase activity (Lohe et al., 1997). The distance between the aspartic acid residues is variable, while the distance to the glutamic acid is dependent on the transposon family. Some are separated by as little as 34 amino acids, and some up to 154

amino acids. It is believed that this motif plays a role in coordinating metal cations that are involved in nucleophilic attacking groups (Mahillon and Chandler, 1998).

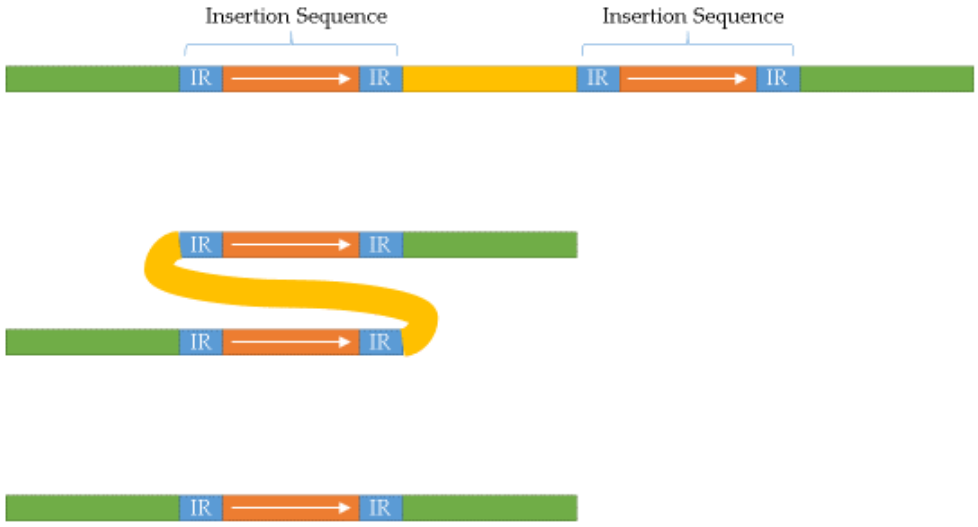
Errors can occur during the excision of IS elements via the DNA replication or repair proteins which can result in mutations. Some DNA from the IS element can remain in the host genome resulting in an insertion. Conversely, the IS element may take some host DNA with it during excision, resulting in a genomic deletion (Kusumoto et al., 2011). IS elements interacting with sequences that share substantial nucleotide identity can lead to larger scale genomic rearrangements. Recombination of two IS elements in the same orientation results in a deletion, and recombination between two IS elements in the opposite orientation would result in an inversion of the intervening sequence (Darmon and Leach, 2014) (see Figure 1).

Regulation of IS Elements in Bacteria

All of these mobile genetic elements allow bacteria to adapt to their changing surroundings, but too much genetic movement can be harmful to the bacteria. Therefore, there are many different ways that mobile genetic elements are regulated. Transposase promoters are generally weak and found in the inverted repeats, which allows for autoregulation of the transposase (Mahillon and Chandler, 1998). Transcriptional repressors or inhibitors can be involved in the regulation of transposase expression (Escoubas et al., 1991). The translation initiation signal may be sequestered by secondary structures formed from neighboring DNA transcripts of mRNA. It is believed that intermediates may not be formed due to impinging transcription, which then inhibits transposition activity (Mahillon and Chandler, 1998).

Many transposons encode antisense RNA opposite of their transposase gene. These have been shown to reduce transposase levels and thus inhibit transposition (Thomason and Storz, 2010; Wagner et al., 2002). Antisense RNAs of transposases are

A



B

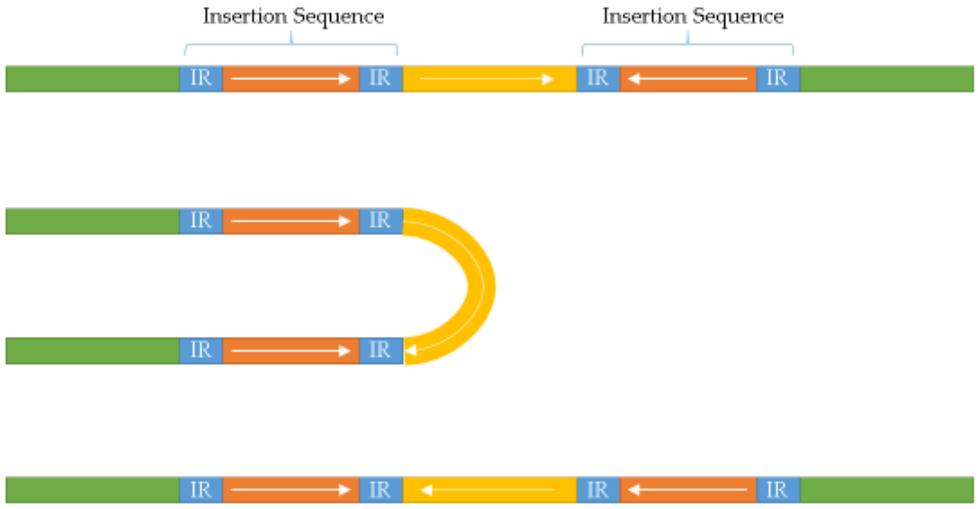


Figure 1: Recombination between two insertion sequence elements. (A) Shows recombination resulting in a deletion of the intervening sequence. (B) Shows recombination resulting in an inversion of the intervening sequence.

present constitutively in bacteria and act as repressors, allowing transposase expression to occur only during certain conditions and repressed under normal circumstances (Thomason and Storz, 2010). This tight regulation of transposase expression provides genome stability.

Programmed translational frameshifting can also occur in transposases. This is usually a -1 frameshift at a sequence similar to AAAAAAG in which the ribosome slips backwards one base upstream. After ribosomal slippage, it then continues translating in a different reading frame. Since lysine is encoded by codons AAA and AAG, lysine-tRNA deprivation can induce frameshifting by causing the ribosome to pause and slip while stalled at this codon (Farabaugh, 1996). The frequency of the frameshifting directly relates to the transposition activity.

The stability of the transposase will also affect transposition activity. Many factors play a role in transposase stability, such as mutations, temperature sensitivity and other factors (Mahillon and Chandler, 1998). Other factors can influence transposase stability as well, such as the Lon protease in *Escherichia coli*. It was found that Lon protease limits the activity of *E. coli* transposase IS*903* while mutant strains lacking the Lon protease gene produced more stable transposases (Derbyshire et al., 1990). Each of these factors could cause the transposase to become unstable and create a temporary or irreversible loss of activity.

Transposases are also more effective if the IS element encodes that transposase, or there is another transposase gene close by in the genome. The *cis* preference for nearby genes is another control mechanism to prevent one transposase from activating transposition in other areas of the genome. Too much transposition can lead to genomic instability. In general, the higher the abundance of genes encoding transposases in the

genome and the more stable the expressed transposase, the more these elements can act in *trans* (Mahillon and Chandler, 1998).

Other bacterial factors also play a role in the regulation of IS elements and a variety of mechanisms may affect transposition. Some host proteins, such as integration host factor (IHF) have binding sites within IS elements which can lead to the direct regulation of expression (Grindley and Joyce, 1980). DNA chaperones can assist with forming the correct secondary or tertiary structures. Other proteins involved with proteolysis can degrade the expressed transposases. High expression levels of transposases can also induce the SOS response in the host (Weinreich et al., 1991). However, it has also been found that constitutive SOS conditions can enhance transposition (Kuan et al., 1991). Finally, Dam DNA methylase can affect transposase expression due to methylation of the site GATC in most transposase promoter regions (Mahillon and Chandler, 1998).

Mobile Genetic Elements Influence on Virulence

IS elements can affect bacterial pathogenicity by altering gene expression, such as in case of antigenic or phase variation, contributing to immune evasion. Insertion of an IS element upstream of a gene may also lead to up-regulation of that gene, as transcription of the IS element can continue past this gene into downstream genes. Some IS elements even contain outward promoter sites near their terminal repeats which can lead to expression of surrounding genes, which can include virulence genes. (Safi et al., 2004). When an IS element disrupts a gene, or its regulatory sequence, those genes can become inactivated. For example, in *Pseudomonas atlantica,* IS492 inactivates extracellular polysaccharide production. However, when IS492 is excised, extracellular polysaccharides are able to be produced, contributing to biofilm formation (Perkins-Balding et al., 1999).

IS elements are often very abundant in the genomes of highly virulent pathogens, including the select agents *Francisella tularensis*, *Burkholderia mallei*, and *Yersinia pestis*. These species have all acquired virulence factors through lateral gene transfer and contain many mobile genetic elements. They have also undergone genome reduction compared to their ancestors that is possibly mediated by insertion sequences (Keim and Wagner, 2009).

Yersinia pestis contains many IS elements, accounting for 3.7% of the genome, which is about tenfold higher than *Y. pseudotuberculosis*. *Y. pestis* also contains rearrangements of local collinear blocks that are surrounded by IS elements (Parkhill et al., 2001). *Burkholderia mallei* has 171 IS elements, accounting for 3.1% of its genome, while *B. pseudomallei* and *B. thailandensis* only contain 42 and 46, respectively (Nierman et al., 2004). *F. tularensis* also contains many IS elements and local collinear block rearrangements that are surrounded by IS elements (Larson et al., 2015). We hypothesize that IS elements in *F. tularensis* contribute to differential gene expression among the various subspecies and subtypes. Therefore, IS element associated gene expression in *F. tularensis* was the focus of this investigation and is discussed in more detail below.

Francisella tularensis Classification and Historical Perspective

The genus *Francisella* currently contains six species, specifically *F. halioticida, F. hispaniensis, F. noatunensis, F. persica, F. philomiragia*, and *F. tularensis*. *F. tularensis* is a gram negative coccobacillus and the causative agent of the disease tularemia. There have been over 250 species that are known to be infected by *F. tularensis*, including mammals, birds, reptiles, fish, and invertebrates (Marner and Addison, 2001). *F. tularensis* is present across the Northern Hemisphere, and is a select agent due to its high infectivity, easy dissemination (aerosolization), and capacity to cause disease (Dennis et al., 2001). As little as 10 organisms can cause disease and can persist for long periods of time in

moist environments. *Francisella* was proposed to be the causative agent of many ancient plagues, including the Hittite plague, Philistine plague, and an Egyptian plague (Trevisanato, 2004; Trevisanato, 2007a; Trevisanato, 2007b). It was studied by the Japanese during World War II in Manchuria (Dennis et al., 2001). Ken Alibek, a former director of the Soviet biological weapons organization, suggested that a *F. tularensis* outbreak on the German-Soviet front shortly before the Battle of Stalingrad in 1942 may have been intentional (Geissler, 2005). It has also been studied, weaponized, and stockpiled by Japan, the United States, and Russia. Those stockpiles have since been presumably destroyed.

F. tularensis is comprised of four subspecies, *F. tularensis* subsp. *tularensis* (type A, further divided into type A.I and type A.II, with A.I being more virulent than A.II), *F. tularensis* subsp. *holarctica* (type B), *F. tularensis* subsp. *mediasiatica,* and finally, *F. tularensis* subsp. *novicida*. The classification of *F. tularensis* subsp. *novicida* is in dispute. Originally identified as *F. novicida*, this organism was reclassified as *F. tularensis* subsp. *novicida* in 2010 (Huber et al., 2010), however, *F. novicida* is still on the Approved List of Bacterial Names (Skerman et al., 1980) The reclassification is currently in contention (Busse et al., 2010; Johansson et al., 2010). Further complicating the issue is the fact that *F. tularensis* (*F. t.*) subsp. *novicida* is published in different journals under different names (*F. tularensis, F. tularensis* subsp. *novicida*, *F. novicida*, etc.) (Kingry and Petersen, 2014). Nevertheless, the International Journal of Systematic and Evolutionary Microbiology is the official journal for naming prokaryotes, therefore that nomenclature that will be used herein is *F. t. novicida*.

F. t. tularensis is only found in North America. Type A.I is generally found in eastern and central United States, while type A.II is found in the western United States. This is thought to be in relation to the different vectors that each subspecies inhabits. *F. t.*

holarctica is found throughout Europe, Japan, and North America. In Russia and Asia the prevalent species is *F. t. mediasiatica* (Farlow et al., 2005). The last subspecies, *F. t. novicida* is found in North America, though it has been recently isolated in Australia and Thailand (Carvalho et al., 2014).

F. t. *tularensis* generally has a terrestrial lifecycle, while *F. t. holarctica* has a more aquatic lifestyle. Arthropod vectors are important for transmitting *F. tularensis* between mammalian hosts (Ellis et al., 2002). There are many different kinds of ticks, including *Dermacentor, Ixodes,* and *Amblyomma*, along with mosquitoes (*Aedes, Culex,* and *Anopheles*), and biting flies (*Chrysops discalis*) that can carry *F. tularensis* (Ellis et al., 2002; Farlow et al., 2005). The geographical distribution of *F. t. tularensis* type A.I and type A. II in the United States corresponds to the different species of ticks that carry each type. *Dermacentor variabilis* and *Amblyomma americanum,* which carry type A.I, are in the central and eastern United States and in California. While *D. andersoni*, which carries type A.II, is found in the western United States (Farlow et al., 2005). There are numerous mammalian hosts that *F. tularensis* has been found in, including humans, rabbits, hares, cats, skunks, squirrels, muskrats, and voles. There is also evidence that amoebas may become infected with *F. tularensis*, allowing the bacteria to persist in water (Carvalho et al., 2014). *F. tularensis* is known to persist is water and mud for 14 weeks, oats for 4 months, and straw for 6 months (Feldman et al., 2001).

There have been several outbreaks of *F. tularensis* associated with different sources. There was an outbreak in Spain associated with crayfish, an outbreak in South Dakota associated with ticks from dogs, and outbreaks in Turkey and Bulgaria associated with contaminated water (Foley and Nieto, 2010). Given that tularemia is also known as “rabbit fever,” it is not surprising that rabbits and hares have been associated with multiple human outbreaks, one well known outbreak occurred on Martha’s Vineyard. There were

15 patients with tularemia, of which 11 had pneumonic tularemia, 2 had ulceroglandular, and the remaining 2 had typhoidal tularemia. One patient died from pneumonic tularemia due to a delay seeking medical attention after becoming ill. Through an epidemiological study, it was found that lawn mowing and brush cutting were associated with this outbreak. It is believed that directly mowing over a carcass or animal excrement aerosolized the bacteria and which was then inhaled by the people (Feldman et al., 2001).

There are different manifestations of a *F. tularensis* infection ranging from ulceroglandular, pneumonic, oculoglandular, oropharyngeal, and typhoidal. If untreated, infections can become systemic. The average incubation period is 3-5 days, but can be as long as 20 days (Carvalho et al., 2014). General symptoms include fever, chills, myalgia, and headaches. Prior to antibiotics, fatality rates were 30-60% depending on the type of infection (Ellis et al., 2002).

Vector-borne transmission, or handling infected animals can result in an ulceroglandular infection (Sjöstedt, 2007). Ulceroglandular is the most common form of tularemia. An ulcer will develop at the site of infection and can continue for months and possibly leave a scar. A fever and lymphadenopathy can develop, and can progress to pneumonia or septicemia (Carvalho et al., 2014; Foley and Nieto, 2010).

Pneumonic tularemia often results from aerosols from infected animals or feces. Pneumonic tularemia can also become a complication from other forms of tularemia. Symptoms include coughing, chest pains, dyspnea, and fever. A further complication can be acute respiratory distress syndrome. This form is also the most common laboratory acquired infection of tularemia (Foley and Nieto, 2010).

Ocularglandular infections arise from direct inoculation of the eye and are rare. Oropharyngeal (gastrointestinal) tularemia is the result of ingesting contaminated food or water. There can be a sore throat, abdominal pain, vomiting, diarrhea, and GI bleeding

(Foley and Nieto, 2010). Typhoidal tularemia occurs when the infection is systemic, but the route of infection is unclear. Typhoidal tularemia is septicemia without lymphadenopathy and often presents with pneumonic tularemia. Oropharyngeal and typhoidal tularemia may cause the patients to go into shock (Ellis et al., 2002).

F. tularensis is resistant to beta-lactam antibiotics. Aminoglycosides are used in the most serious cases, but complications include ototoxicity and nephrotoxicity, leading to a decline in use. However, gentamicin has been used to treat pneumonic tularemia (Carvalho et al., 2014). Chloramphenicol is effective against *F. tularensis*, but is rarely used due to its side effects. Fluoroquinolones (specifically ciprofloxacin) are the most widely used for uncomplicated tularemia. Treatment with these drugs must be continued for at least 14 days to clear the infection (Foley and Nieto, 2010).

Francisella Pathogenicity

F. tularensis genome contains one or two pathogenicity islands (FPI) (one copy in *F. t*. *novicida* and two copies in *F. t. tularensis*, *F. t*. *holarctica*, and *F. t*. *mediasiatica*). The FPI has a cluster of 13-19 ORFs that have a lower G+C content then the rest of the genome, suggesting it was acquired via horizontal gene transfer (Broms et al., 2010). It is believed that the duplicate FPI in virulent *F. tularensis* strains was due to the insertion of an IS element, IS*Ftu1,* adjacent to the FPI which allowed for non-reciprocal recombination, whereby the donor bacteria provided the additional copy of the FPI to a recipient bacteria without the donor receiving any additional DNA (Larsson et al., 2009). Some studies have suggested that the second FPI is functionally redundant, since targeted deletions/mutations in one copy do not affect virulence (Golovliov et al., 2003; Kawula et al., 2004; Qin and Mann, 2006); however, other studies suggested that both copies of some genes, such as *iglA, iglB, iglC,* and *iglD*, are required for full virulence (Maier et al., 2007; Su et al., 2007).

There is evidence that genes within the FPI encode for a type 6 secretion system (T6SS) in *F. tularensis*. T6SSs are multi-subunit complexes that transport molecules from the bacteria cytoplasm into a host cell (Broms et al., 2010). A T6SS was first identified in *Vibrio cholera* (Das and Chaudhuri, 2003) and has now been found in many other bacterial species. Most T6SS contain homologs to *Vibrio cholera* IcmF, DotU, ClpV, VipA, VipB, VgrG, and Hcp proteins (Boyer et al., 2009). IcmF and DotU are membrane proteins that are part of the secretion apparatus and required for a functional T6SS (Broms et al., 2010; Ma et al., 2009). ClpV is an ATPase that is involved with secretion of T6SS substrates by forming a hexameric complex at the inner membrane (Schlieker et al., 2005; Zheng and Leung, 2007). VipA and VipB are a bicistronic locus that forms a tubule that resembles the T4 bacteriophage tail sheath (Leiman et al., 2009). VgrG likely forms a trimer that has the ability to puncture bacterial and eukaryotic membranes, and some VgrG proteins have C-terminal extensions that are able to interfere with cellular functions in the host cell (Pukatzki et al., 2007). Finally, Hcp is the major secreted protein of the T6SS which forms hexameric donut-shaped rings. These rings stack to form tubules that enable delivery of macromolecules (Mougous et al., 2006).

The suggested model for the function of a T6SS involves the formation of a pore complex in the bacterial plasma membrane in which Hcp and VgrG are secreted through, and ClpV is believed to localize to the membrane pore complex and energize the process (assembly, disassembly, and secretion) (Bönemann et al., 2010; Broms et al., 2010; Filloux, 2009). Hcp assembles into a tube structure that is capped by the VgrG trimer and sheathed with the VipA-VipB tubule. Hcp then elongates while VipA-VipB contracts, allowing VgrG to puncture the bacterial membrane and possibly a host cell membrane. VgrG detaches to allow Hcp and other components to be delivered into the host cell (Broms et al., 2010; Pukatzki et al., 2009).

There are homologs for a T6SS in the FPI of *F. tularensis.* The IcmF homolog in *F. tularensis* is PdpB, and DotU is also found in the FPI. Both PdpB and DotU have predicted transmembrane domains (Broms et al., 2010). IglA and IglB in *F. tularensis* are homologs of VipA and VipB. There is also evidence that the *F. tularensis* FPI contains a smaller VgrG homolog than any other known VgrG, and lacks the C-terminal domain. Although *F. tularensis* IglF was proposed to be a ClpV homolog, it lacks the ATPase domain of other ClpV proteins (Broms et al., 2010). PdpE has been proposed to be a possible homolog of the Hcp protein, but the homology is faint and there is no evidence that it performs a similar function (Barker et al., 2009). Therefore, the non-canonical gene composition of the *F. tularensis* T6SS suggests a unique mode of action.

The FPI encodes *iglA, iglB, iglC, iglD,* and *pdpA,* which are essential for intracellular survival (Chong et al., 2008). In *F. t*. *holarctica*, the *anmK* gene and most of *pdpD* gene are missing. This may have some effect on the virulence differences between type A and type B strains (Nano et al., 2004). IglC and other FPI genes play a crucial role in phagosome escape and avoiding lysosomal fusion (Santic et al., 2010). FPI genes are expressed throughout cytosolic replication. It is unknown if they play a role in the late stages of infection or if this is in preparation for infecting new cells (“ready and armed” strategy) once their current host is lysed (Wehrly et al., 2009).

FPI gene regulation has been shown to be controlled by at least six proteins, MglA, SspA, FevR, MigR, Hfq, and PmrA (Broms et al., 2010). Some of these regulatory proteins are also involved with regulation outside of the FPI. The most studied of these regulatory proteins has been MglA. For instance, MglA and SspA are known to interact and form a heterodimer. This heterodimer will bind to RNA polymerase and induce transcription of certain genes. FevR will also interact with this heterodimer. Phosphorylated PmrA may bind to FPI gene promoters and recruit the MglA, SspA, and FevR to initiate transcription

(Broms et al., 2010). MigR is thought to influence FPI gene expression through indirect interaction with FevR (Buchan et al., 2009). Hfq acts as a repressor, while all the other known regulators act on promoters (Meibom et al., 2009).

F. tularensis also encodes a putative type IV pili. Type IV pili are flexible filamentous appendages that have been known to facilitate attachment to host cells, DNA update, and biofilm formation in other bacteria (McLendon et al., 2006). PilA is the major pilin that is processed by PilD and translocated across the inner membrane. Once assembled into a pilus fiber, it is then secreted through the secretion pore PilQ which is facilitated by the PilB ATPase. Disassembly and retraction are facilitated by PilT (Salomonsson et al., 2011). It has been found in *F. tularensis* that PilA is required for virulence in type A and type B strains (Forslund et al., 2010). PilC and PilQ are involved with secretion and assembly of type IV pili in *F. tularensis,* and have also been found to contribute to virulence (Forslund et al., 2010). Type B strain Live Vaccine Strain (LVS) has been shown to express type IV pili-like structures on its surface and mutations in *pilB* and *pilT* led to a loss of those surface fibers (Gil et al., 2004). However, type IV pili are not required for intracellular survival or replication in type A and B strains, although *pilA* is required for intracellular growth of *F. tularensis* subsp. *novicida* (Forslund et al., 2006).

Virulent *F. tularensis* strains only induce a weak proinflammatory response by macrophages, which is in part due to its unique lipopolysaccharide (LPS) that presumably does not interact with the toll-like receptor 4 (TLR4) (Gunn and Ernst, 2007). The O-antigen of LPS of *F. tularensis* plays a role in avoidance death when first taken up by macrophages (Carvalho et al., 2014). LPS differs from that in other species due the absence of a phosphate at the 4’ position on the non-reducing glucosamine backbone, and containing longer acyl side chains of 16-18 carbons compared to 12-14 in other

bacteria (Gunn and Ernst, 2007). Therefore, *F. tularensis* LPS has low toxicity and induces lower levels of cytokines (Ellis et al., 2002).

The capsule of *F. tularensis* is composed of an O-antigen identical to the O-antigen of LPS, but does not contain other LPS components (Rowe and Huntley, 2015). The capsule is important for replication in the cytosol following phagosomal escape; without the capsule, *F. tularensis* is exposed to host intracellular detection pathways that can trigger apoptosis (Lindemann et al., 2011). Studies have demonstrated that increased capsule production in LVS lead to a more virulent strain in a mouse model, while an acapsular strain was attenuated (Bandara et al., 2011; Cherwonogrodzky et al., 1994; Sandström et al., 1988). However, the role of the capsule in serum resistance is not completely understood at this time (Rowe and Huntley, 2015).

F. tularensis can infect and replicate inside macrophages. *F. tularensis* is taken up by host macrophages via induction of spacious asymmetric pseudopod loops that is mediated by complement factor C3 and complement receptor 3 (CR3), and possibly other complement receptors such as mannose receptor and scavenger receptor class A, into a *Francisella*-containing phagosome (FCP) (Clemens et al., 2005; Santic et al., 2010). The phagosome does not mature further and becomes only moderately acidic, which is then followed by bacterial escape into the cytosol (McLendon et al., 2006). Interestingly, *F. tularensis* can use acidification dependent and independent mechanisms to escape into the cytosol, but escape is much less efficient without acidification of the phagosome (Chong et al., 2008). Nevertheless, *F. tularensis* is able to escape the FCP before the lysosome fuses and elicits no proinflammatory response, although exactly how *F. tularensis* is able to do this is not fully elucidated yet (Asare and Abu Kwaik, 2010; Clemens et al., 2004).

F. tularensis then proceeds to multiply in the cytosol and eventually causes apoptosis in the host cell (McLendon et al., 2006). *F. tularensis* is able to downregulate TLR 1, 4, 5, 6, 7, and 8 and upregulate TLR2 (Santic et al., 2010). The down regulation of TLRs weakens the host inflammatory response. The innate immune system is suppressed and the profile of chemokines and cytokines that are secreted differ from other bacterial infections. Interferon (IFN)-γ is also suppressed by *F. tularensis* through the upregulation of SOCS3. INF-γ triggers nitric oxidase synthase (iNOS) and limits phagosomal escape (McLendon et al., 2006). TLR2 and MyD88 co-localize with *F. tularensis* in infected macrophages and induce the host immune response (Broms et al., 2010). Signaling from TLR2 and MyD88 is essential for expression of IL-1β and INF-β cytokines, which are critical for a proinflammatory response by the infected host cell (Broms et al., 2010; Carvalho et al., 2014). In order for an infection to be cleared, T-lymphocytes $CD4^+$ and $CD8^+$ cells are needed (Ellis et al., 2002).

Francisella and Insertion Sequence Elements

F. tularensis contains several IS elements, with IS*Ftu1* being the most abundant, with the exception of *F. t. novicida* (see Table 1). IS*Ftu1* belongs to the IS*630* Tc-1 mariner family. These transposases recognize a TA sequence for insertion sites, but the surrounding bases play a role in frequency of insertion. The IS*630* target sequence is thought to be 5'-CTAG-3' and creates a duplicate TA at the insertion site (Mahillon and Chandler, 1998). Members of this family contain the DDE (or DDD) motif discussed earlier, which is essential for catalytic activity (Shao and Tu, 2001). These transposases utilize cut-and-paste transposition. They are excised through staggered double-stranded DNA breaks that leave behind 2 bp inverted repeats, which are sealed through cellular repair processes (Plasterk et al., 1999). IS*Ftu1* does contain a slippery heptamer (AAAAAAG) for the ribosome to frameshift -1 to translate the entire transposase. IS elements are not

the only genes that require programmed translational frameshifting for protein production; the *dnaX* gene in *F. tularensis* and other bacterial species (e.g., *E. coli*) encodes DNA polymerase III subunits gamma and tau by utilizing -1 programmed translational frameshifting (Chen et al., 2014; Kodaira et al., 1983).

The amount of IS elements also indicates the genome is undergoing a reduction process and adaptation to an intracellular lifestyle. Many pathogens contain a large number of IS elements and pseudogenes. This follows the "mutational hazard hypothesis: severe reductions in population size result in less effective selection, which promotes the accumulation of non-functional and slightly deleterious sequences." (Bobay and Ochman, 2017). During genome reduction, many bacteria loose genes related to metabolic pathways (Casadevall, 2008), which is the case with *F. tularensis*. The study of IS elements has provided clues to the evolution of *F. tularensis*. As shown in Table 1, *F. tularensis* subsp. *novicida* does not contain as many IS elements as the other *F. tularensis* subspecies and the genome has remained more stable, indicating early divergence (Larsson et al., 2009). *F. tularensis* subsp. *holarctica* emerged more recently and consequently is more clonal (Larson et al., 2011). *F. tularensis* subsp. *novicida* also maintains more metabolic competence than *F. tularensis* type A and B, which further indicates that the virulent species is adapting to a niche restriction as have other pathogens (Larsson et al., 2009).

IS elements in *F. tularensis* have even been used to distinguish the different subspecies. A differential insertion sequence amplification (DISA) PCR assay was designed based on inversions and rearrangements presumably mediated by insertion sequences (Larson et al., 2011). Two regions were identified that gave different banding patterns in a gel after PCR amplification for type A.I, A.II, and B. One region also distinguished a different band for *F. t. novicida*. These results corresponded with subtyping

done by pulse field gel electrophoresis (PFGE). A size difference between amplicons was based on a deletion in the IS*Ftu1* gene, which could indicate IS element inactivation and provide further evidence for genomic reduction. This assay was also successfully used to subtype or identify wild type *Francisella* strains in clinical and environmental samples (Larson et al., 2011).

Insertion sequences have been suggested to play a role in the virulence of *F. tularensis* through the duplication of the FPI (Larson et al., 2015). Another gene that is flanked by IS*Ftu1* encodes a putative glycosyl transferase and is duplicated in the genome of virulent *F. tularensis* strains, but not the less virulent *F. tularensis* subsp. *novicida* (Larsson et al., 2009), and therefore, may contribute to pathogenicity. IS elements certainly contributed to the apparent genomic rearrangements observed within the chromosome of the different *F. tularensis* subspecies and subtypes, affecting the phenotype and virulence, such as the case for the virulent *F. tularensis* strains versus the avirulent *F. tularensis* subsp. *novicida* strains.

IS elements can also affect the virulence of *F. tularensis* by regulating nearby genes. For example, a previous study demonstrated that IS elements are upregulated in the presence of spermine (Carlson et al., 2009). Spermine is a polyamine produced by eukaryotes that is involved in DNA synthesis, transcription, and translation. Carlson and associates infected macrophages with *F. tularensis* LVS and Schu S4 that had been previously grown in media with either the presence or absence of spermine. There was an increased cytokine response from those cultures grown without spermine compared to those with spermine; however, the mechanism contributing to this outcome is unknown. The transcriptome was also examined using custom designed microarrays in this study. This analysis showed the upregulation of transposase genes along with downstream genes, and there was variation in the genes that were upregulated, which corresponded

to the different locations of the IS elements within the *F. tularensis* LVS and Schu S4 genomes (Carlson et al., 2009). These studies provide evidence that when *F. tularensis* infects a host, the host environment can induce the expression of IS elements, altering gene expression and virulence. Therefore, we hypothesize that *F. tularensis* IS elements contribute to differential gene expression that is observed among the various subspecies and subtypes.

Table 1: IS elements (and remnants) in *Francisella*. Numbers of IS elements (and remnants) in *F. tularensis* (*F. t.*) subsp. *tularensis* subtype A.I reference strain Schu S4, *F. t.* subsp. *tularensis* subtype A.II prototype strain WY96-3418, *F. t.* subsp. *holarctica* type B attenuated strain LVS, *F. t.* subsp. *mediasiatica* strain FSC147, and avirulent *F. t.* subsp. *novicida* strain U112, as well as *F. philomiragia* ATCC 25015.

IS Elements	*F. t.* subsp. *tularensis* subtype A.I (Schu S4)	*F. t.* subsp. *tularensis* subtype A.II (WY96-3418)	*F. t.* subsp. *holarctica* type B (LVS)	*F. t.* subsp. *mediasiatica* (FSC147)	*F. t.* subsp. *novicida* (U112)	*F. philomiragia* (ATCC 25015)
IS*Ftu1* (IS630 family)	47 (3)	48 (8)	51 (8)	56 (1)	1	0
IS*Ftu2* (IS5 family)	13 (3)	18 (18)	41	14 (3)	18	0
IS*Ftu3* (ISNCY family, ISHpal-IS1016)	2 (1)	1 (3)	2 (1)	2	4	0
IS*Ftu4* (IS982 family)	1	1	1	1	1	0
IS*Ftu5* (IS4 family)	1	1	3	1	0	0
IS*Ftu6* (IS1595 family)	1 (1)	1 (1)	1 (1)	2 (1)	2	0
ISSod13 (IS3 family)	1	1	1	1	0	0
TOTAL	66 (8)	71 (30)	100 (10)	77 (5)	26	0

MATERIALS AND METHODS

Bacterial Strains and Culturing

The *F. tularensis* strains used in this study included Schu S4, NE061598, WY-00W4114, LVS, and U112. Select agent *F. tularensis* strains were transferred to the University of Nebraska Medical Center in Omaha following the requirements of the Select Agent Program as outlined in the Animal and Plant Health Inspection Service/Centers for Disease Control and Prevention (CDC) Form 2, Guidance Document for Request to Transfer Select Agents and Toxins (Centers for Disease Control and Prevention (CDC) and Department of Health and Human Services (HHS), 2012). Manipulation of viable culture material was performed by authorized individuals within a biosafety level 3 (BSL-3) laboratory certified for select agent work by the United States Department of Health and Human Services using laboratory biosafety criteria, according to requirements of the Federal Select Agent Program. All *F. tularensis* strains were initially grown on Chocolate agar plates (Remel, catalog number R01302) from a master stock and incubated at 37ºC with 5% CO_2 for three days before processing. All RNA derived from select agent strains of *F. tularensis* were confirmed to be sterile before removal from the BSL-3 facility by plating 10% of the isolated RNA on Chocolate agar plates and incubating the plates at 37ºC with 5% CO_2 for not less than five days.

For the *F. tularensis* growth curves, strains were grown in 100 mL of supplemented Brain Heart Infusion (BHI) in unbaffled 500 mL flasks with 5% CO_2 and shaking at 200 rpm at 37ºC. In more detail, BHI was prepared as recommended by the manufacturer (BD Biosciences catalog number 211059) and supplemented with cysteine, β-NAD, hemin, histidine, NaOH, and glucose, as previously described (Mc Gann et al., 2010). The

inoculum for the *F. tularensis* cultures was prepared by aseptically transferring a lawn of bacterial growth into the supplemented BHI and then inoculating 100 mL of supplemented BHI to obtain an initial OD_{600} of 0.05. Thereafter, the absorbance at 600 nm was obtained and recorded every 4 h. Cultures were pelleted at the desired time points and processed for RNA isolation. *F. tularensis* Schu S4, NE061598, and W4114 cultures were at early, mid, and late log growth phase at 12 h, 20 h, and 28 h post-inoculation; 40 h and 64 h post-inoculation correspond to early and late stationary phase, respectively. For select agent *F. tularensis* WY96, 28 h, 40 h, 48 h, and 60 h corresponded to early, mid, and late log, and stationary phase, respectively. The growth curves for *F. tularensis* LVS and U112 were obtained by Eric Miller, using the same medium and growth conditions as was used for the select agent *F. tularensis* strains. The time points at 4 h, 8 h, 12 h, and 20 h corresponded to early, mid, and late log phase and stationary phase, respectively.

SNAP was added to the *F. tularensis* NE061598 culture during early log growth phase, which was approximately 11.5 h post-inoculation. The SNAP reagent was prepared by dissolving 0.004 g of SNAP (Sigma catalog number N3398) in 40 µL DMSO, and then 20 µL of this 0.45 M SNAP solution was added to the 100 mL culture, in order to obtain a final SNAP concentration of 90 µM. The addition of spermine to *F. tularensis* NE061598 culture in early log growth phase was also performed. The spermine reagent was prepared by adding 0.04 g spermine (Sigma catalog number S3256) to 10 mL of sterile water, subsequent filter sterilization using a 0.22 µm filter, and then 1 mL of this 20 mM spermine solution was added to a 100 mL culture of NE061598, in order to obtain a final concentration of 200 µM. The growth of *F. tularensis* NE061598 was then monitored and RNA was isolated after 30 minutes of treatment with either SNAP or spermine, as described herein for the untreated *F. tularensis* cultures.

RNA Isolation

F. tularensis RNA was isolated at time points of interest using RNeasy Mini Kit (Qiagen catalog number 74104), as recommended by the manufacturer. A total of 1 x 10^9 cells were pelleted for subsequent RNA extraction at each time point and supernatants were discarded. RLT buffer was prepared by adding 120 µL β-mercapto-ethanol to 12 mL of RLT buffer. Cell pellets were then resuspended in 900 µL RLT buffer and transferred to an appropriate tube containing 0.1 mm diameter glass beads (BioSpec catalog number 11079101). Samples were then bead beat at a speed setting of 6.0 for 23 sec and centrifuged at 13,000 rpm for 2 min. Approximately 700 µL of the supernatant was transferred to a new tube and then 500 µL of 100% ethanol was added. Half of the mixture was transferred to a spin column and centrifuged at full speed for 1 min, flow thru was discarded, and then the remaining sample was added and centrifuged again. The spin column was washed with 350 µL RW1 buffer and centrifuged again. A DNase mixture comprised of 70 µL RDD buffer and 10 µL DNase was added to the spin column and incubated at room temperature for 20 min. Spin columns were then washed again with 350 µL RW1 and centrifuged, followed by two washed with 500 µl RPE and centrifugation. The spin column was then transferred to a new tube and 50 µL of RNase-free water was added and incubated at room temperature for 1 min followed by centrifugation. RNA was stored at -80ºC until sterility was confirmed for the select agent *F. tularensis* strains, as previously described, and/or needed for further processing.

The integrity of the isolated RNA was checked by boiling approximately 100 ng of RNA at 75ºC for 10 min, adding loading buffer, running samples in a 1% agarose gel containing ethidium bromide for visualization of the nucleic acid (0.5 g agarose, 50 mL 1x TAE buffer, 1 µL ethidium bromide [10 mg/mL]). Gels were ran at 90 constant volts (cV) for 1 h. Gel pictures were taken with UV setting of AlphaImager from ProteinSimple.

Nucleic Acid Quantitation

To determine the concentration of RNA or DNA, a Qubit fluorometer was used along with the Qubit dsDNA BR Assay Kit (Invitrogen, catalog number Q32850), Qubit dsDNA HS Assay Kit (Invitrogen, catalog number Q32851), or Qubit RNA HS Assay Kit (Invitrogen, Q32852). A master mix of Qubit buffer and Qubit reagent was made by adding 199 µL x n (where n equals the number of samples, plus two standards, plus one) Qubit buffer and 1 µL x n Qubit reagent. For the high and low concentration standards, the Qubit master mix was aliquoted (190 µL) into two Qubit assay tubes (Invitrogen, catalog number Q32856) for each standard, and then 10 µL of the appropriate standard #1 was added to one tube and 10 µL of the associated standard #2 was added to the other tube. For each sample, 190-199 µL Qubit master mix was added to a Qubit assay tube and 10-1 µL of the sample was added for a final volume of 200 µL. All standards and samples were vortexed for 2-3 sec and incubated at room temperature for 2 min. The two standards were used to calibrate the Qubit and then the sample concentrations were obtained.

DIG Labeled Probe Synthesis

To prepare the IS*Ftu1* probe for northern blot analysis, PCR amplification of the IS*Ftu1* gene was performed using 100 ng of genomic DNA from *F. tularensis* NE061598 as the template along with the primers F206 and R201 (Table 2) and the Platinum *Taq* DNA Polymerase High Fidelity Kit (Invitrogen catalog number 10966018), as recommended by the manufacturer. The PCR cycling conditions were as follows: 2 min denaturation at 95^{o} C, followed by 35 cycles of 95^{o}C for 30 sec, 53^{o}C for 30 sec and a 1 min extension at 72^{o}C with a final 7 min extension for 7 min. The resulting 144 bp IS*Ftu1* amplicon was purified using Zymo's DNA Clean and Concentrator Kit, according to the manufacturer's recommendations.

Next, the purified IS*Ftu1* PCR product was DIG labeled using the same primers and PCR cycling conditions described above along with the Roche PCR DIG Probe Synthesis Kit (catalog number 11 636 090 910), as recommended by the manufacturer. For the labeling reactions, the 144 bp IS*Ftu1* PCR product was diluted to 20 pg/µL and used in one unlabeled reaction (control) and one DIG labeled reaction. The DIG labeled probe was then checked for quality and labeling efficiency by running an aliquot of both the labeled probe and unlabeled product in a 1% agarose gel.

Northern Blot Analysis

The *F. tularensis* RNA was fractionated in a 50 mL 1% agarose gel prepared with 5 mL 10x MOPS buffer (Lonza Walkersville catalog number 50876) and 2.8 mL formaldehyde for subsequent northern blot analysis. RNA (5 µg) from mid exponential growth phase *F. tularensis* Schu S4 and NE061598 was mixed with 20 µL sample buffer that contained 250 µL formamide, 90 µL formaldehyde, 50 µL 10x MOPS buffer, and 0.5 µL ethidium bromide. The RNA with sample buffer was then boiled for 10 mins at 75^{o}C and cooled on ice. Prior to adding the RNA to the 1% agarose gel wells, 5 µL of 10x Blue Juice loading buffer (Invitrogen catalog number 10816015) was added to each sample. The RNA was fractionated in the agarose gel using 1x MOPS buffer and ran at 70 cV until the first dye front in the loading buffer migrated to approximately 1 inch from the bottom of the gel. The 16S and 23S ribosomal RNA (rRNA) bands were visualized using a UV transilluminator to check RNA integrity. The agarose gel was then washed twice in sterile water for 10 min per wash. A positively charged nylon membrane (Roche catalog number 11 209 272 001) was cut slightly larger than the gel and washed twice in sterile water for 10 min per wash. Both the agarose gel with the *F. tularensis* RNA and the nylon membrane were equilibrated in 20x SSC buffer (0.3 M Na_3 Citrate, 3 M NaCl, pH 7.0) for 10 min.

The transfer of *F. tularensis* RNA from the gel onto the nylon membrane was accomplished by capillary action or wicking. For this setup, a support was placed in a glass pan and Whatman paper (Fisher catalog number 05-714-4) was hung over the support with the opposite sides of the paper touching the bottom of the pan. The glass pan was then filled with 20x SSC buffer. Next, the gel with the *F. tularensis* RNA was placed face down on the Whatman paper with the nylon membrane on top of the gel. Two more pieces of Whatman paper were placed on top of the gel, after being soaked in 20x SSC buffer. To prevent evaporation of the buffer, Saran wrap was used to cover the entire apparatus with a hole cut out over the membrane. Paper towels were then stacked on top and topped off with a weight. This setup was left overnight at room temperature for the transfer of the *F. tularensis* RNA to the nylon membrane. The following day, the transfer apparatus was disassembled and the nylon membrane containing the transferred *F. tularensis* RNA was crosslinked with UV Stratalinker 1800 twice, using the auto-crosslink setting (preset exposure of 1200 microjoules for 25-50 seconds), turning the membrane 90 degrees for the second crosslinking. UV shadowing was used to visualize and confirm the transfer of the 16S and 23S rRNA onto the nylon membrane, and then the location of these RNA bands were marked on the membrane with a pencil.

For the northern blot hybridization, the membrane containing the transferred *F. tularensis* RNA was placed into a glass hybridization bottle and 10-15 mL of DIG Easy Hyb Buffer (Roche catalog number 11 603 558 001) was added. The membrane and buffer were incubated at 55^{o}C and rotated at a medium speed in a Hybaid oven for 2 h. The DIG labeled IS*Ftu1* probe was thawed in a 55^{o}C water bath 1 h prior to use, and then 50 µL of the probe was added to 10 mL DIG Easy Hyb Buffer and incubated in a 65^{o}C water bath for 10 min. Next, the DIG Easy Hyb buffer in the hybridization bottle was discarded and the DIG labeled IS*Ftu1* probe in the 10 mL buffer was added to the hybridization bottle

containing the membrane with *F. tularensis*. The hybridization mixture was incubated overnight at 55^{o}C in the Hybaid oven while being mixed at a medium speed.

After hybridization, the RNA blot was washed twice with 2x SSC buffer containing 0.1% SDS in the Hybaid oven for 10 min at room temperature. Two additional washes were performed with 0.1x SSC buffer containing the 0.1% SDS in the hybridization oven at 60^{o}C for 15 min. The blot was then transferred to a clean container and washed with 1x Malic acid for 2 min at room temperature, followed by blocking with 50 mL of blocking solution (10% Blocking Solution and 90% 1x Malic acid, Roche catalog number 11 585 762 001) for 2 h at room temperature. After the incubation, the blocking solution was discarded and 1.5 µL of Anti-Digoxigenin AP (Roche catalog number 11 093 274 910) was added to 15 mL of new blocking solution (10% Blocking Solution and 90% 1x Malic acid) for the detection process. The mixture was then added to the container containing the blot. The blot was incubated for 30 min at room temperature with the Anti-Digoxigenin AP mixture and then washed twice with 1x Malic acid for 15 min per each wash at room temperature. Next, the blot was equilibrated in 20 mL of 1x detection buffer (18 mL 1x Malic acid and 2 mL 10x Detection Buffer, Roche catalog number 11 585 762 001) at room temperature for 3 min. A CDP-Star Detection Reagent solution (Roche catalog number 11 685 227 001) was prepared by adding 50 µL of this reagent to 5 mL solution containing 4.5 mL 1x Malic acid and 0.5 mL 10x Detection Buffer. The membrane was then transferred to a plastic bag (Kapak SealPak pouches catalog number 402) and the CDP-Star solution was added to the bag. The bag was heat sealed and incubated for 5 min at room temperature. After incubation, the CDP-Star solution was poured out and the bag was resealed. The northern blot hybridization was imaged with a chemiluminescent camera from Cell Biosciences after a 2 min exposure.

DNase Treatment of RNA

For RT-PCR and deep RNA sequencing, the sterile *F. tularensis* RNA samples were DNase treated using Baseline-ZERO Kit (Epicentre catalog number DB0715K), according to manufacturer protocol. For this treatment, 5 µL of 10x Baseline-Zero Reaction Buffer and 2.5 µL of Baseline-ZERO DNase was added to each 45 µL of *F. tularensis* RNA sample. Reactions were incubated at 37^{o}C for 30 min, followed by the addition of 5 µL Baseline-ZERO Stop Solution and incubation at 65^{o}C for 10 min. The DNase-treated RNA samples were then checked for DNA contamination using *Francisella*-specific CR10 C and L primers, as previously described (Larson et al., 2011). If needed, to remove all of the residual genomic DNA in the *F. tularensis* RNA samples, DNase treatment of the *F. tularensis* RNA was repeated until no amplification product was produced with the CR primer pair, but the positive and negative controls gave the expected results.

Ribosomal RNA Depletion of RNA

To reduce *F. tularensis* rRNA in the samples for subsequent deep RNA sequencing, the MICROBExpress Bacterial mRNA Enrichment Kit (Invitrogen catalog number AM1905) was utilized, according to manufacturer's recommendations. Briefly, 5-10 µg of DNase-treated *F. tularensis* RNA was added to 200 µL binding buffer and 4 µL Capture Oligo Mix and incubated at 70^{o}C for 10 min, followed by 37^{o}C for 15 min. Next, 50 µL of Oligo MagBeads, which was needed for each sample, was aliquoted and washed by placing the bead solution on a magnetic stand, removing and discarding the unbound solution, and then resuspending the beads in nuclease-free water. This bead washing process was repeated for a total of three times and after the second and third wash step, the beads were resuspended in Binding Buffer. The washed beads (50 µL per sample) were incubated at 37^{o}C then added to each RNA/Capture Oligo Mix and incubated at 37^{o}C

for 15 min. The RNA/bead mixture was next placed in magnetic stand and the unbound solution containing the rRNA depleted RNA was transferred to sterile tube. To recover the bound transcripts, 100 µL of Wash Solution, pre-warmed to 37°C, was added to the beads. Beads were then resuspended and the mixture was placed on magnetic stand again. Unbound solution was transferred to previous tube containing the rRNA depleted transcripts for subsequent precipitation. The rRNA depleted RNA was precipitated by adding 35 µL 3M sodium acetate and 7 µL glycogen to each sample and then mixed. Ice cold 100% ethanol (2.5 volumes) was added to each sample and again mixed. After incubation of the RNA for approximately 1h at -80°C, the RNA was pelleted by centrifugation at 13,000 rpm for 30 min at 4°C. The supernatant was discarded and each RNA pellet was washed with 1 mL of ice cold 70% ethanol followed by centrifugation at 13,000 rpms for 10 min at 4°C. The washing of the RNA pellets was performed one additional time, as described above. Finally, RNA pellets were air dried for 5 min at room temperature and then resuspended in 25 µL nuclease-free water. To evaluate the effectiveness of the rRNA depletion process, the rRNA depleted RNA was diluted to 100 ng/µL and checked for rRNA content using an Agilent 2100 Bioanalyzer. If needed, a second ribosomal depletion process was performed.

RNA-Seq Library Synthesis

RNA-Seq libraries were prepared according to manufacturer's recommendations using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Illumina catalog number SSV21124). Approximately 50 ng of rRNA depleted RNA was mixed together with 1 µL RNA Fragmentation Solution, 2 µL cDNA Synthesis Primer, 1 µL of Random Primers at 100 ng/µL, and nuclease-free water to obtain a final volume of 13.5 µL for each sample. The sample was incubated at 85°C for 5 min and then cooled on ice. The cDNA Synthesis Master Mix was prepared by adding 0.3 µL of a dNTP mixture that contained 10 mM of

each dNTP, 3 µL of the cDNA Synthesis Premix, and 0.5 µL of Star Script Reverse Transcriptase per each sample. Next, 3.8 µL of the cDNA Synthesis Master Mix was added to each fragmented sample. The sample was mixed and incubated in a Veriti Thermocycler (Applied Biosystems) at 25ºC for 5 min, 42ºC for 20 min, and then 37ºC. Once the sample reached 37ºC, 1 µL of Finishing Solution was added. The sample was mixed and placed back in the thermocycler for 10 min at 37ºC, 95ºC for 3 min, and then 25ºC. Once the sample reached 25ºC, 8 µL Terminal Tagging Solution containing 7.5 µL Terminal Tagging Premix and 0.5 µL DNA Polymerase was added to each sample. The sample was mixed and again placed back in thermocycler. The sample was incubated at 25ºC for 15 min, followed by 95ºC for 3 min, and then held at 4ºC.

Each cDNA sample was purified using the Agencourt AMPure XP System (Beckman Coulter catalog number A63880), as recommended by the manufacturer and described below in more detail. The purified cDNA was next mixed with 25 µL FailSafe PCR PreMix E, 1 µL Forward PCR primer, 1 µL ScriptSeq Index PCR Primer, and 1 µL FailSafe PCR Enzyme, to prepare for barcoding and amplification. Next, the cDNA was denatured at 95ºC for 1 min and then amplified with 15 cycles of 95ºC for 30 sec, 55ºC for 30 sec, and 68ºC for 3 min with a final extension of 68ºC for 7 min. The barcoded and amplified cDNA was again purified with Agencourt AMPure XP System and then the concentration of the purified cDNA was determined using a Qubit fluorimeter, as described above. The cDNA samples to be sequenced were then diluted to a concentration of 200 nm in Elution Buffer, which was prepared by mixing together 1 µL Tween20 and 999 µL Qiagen Elution Buffer. To prepare the cDNA samples for sequencing, 3 µL of each barcoded cDNA sample was combined into one tube for a final loading concentration of 10 pM. The cDNA samples were then ran at The University of Nebraska Next Generation

DNA Sequencing Core on the Illumina HiSeq2000 platform, using the 100 base pair paired-end run parameter.

Purification of cDNA for RNA-Seq

Samples for RNA-Seq were purified using Agencourt AMPure XP System (catalog number A63880). Prior to purification, the AMPure XP beads were warmed to room temperature and 400 µL of 80% ethanol was made for each sample. AMPure XP beads were vortexed just prior to use to obtain a homogenous solution, and then 45 µL of beads were added to each cDNA sample tube for the first purification or 50 µL of beads for the second purification of the cDNA. Each cDNA sample with the added beads was transferred to a new 1.5 mL tube and incubated at room temperature for 15 min. Sample tubes were then placed on a magnetic stand until pelleted, and then the unbound solution was discarded. Beads with bound cDNA were washed with 200 µL 80% ethanol while still on the magnetic stand for 30 sec, and then the unbound solution was discarded. The 80% ethanol wash was repeated and then the purified cDNA samples were air dried on the magnetic stand for 15 min. Beads with bound cDNA were resuspended in 24.5 µL nuclease-free water for the first purification and 20 µL nuclease-free water for the second purification. To elute the purified cDNA from the beads, samples were incubated for 2 min at room temperature, and then placed on the magnetic stand until beads were pelleted. The unbound solution containing the purified cDNA was then transferred to a new 1.5 mL tube for further processing or stored at -80^{o}C until needed.

Reverse transcription-PCR

The reverse transcription-polymerase chain reaction (RT-PCR) for the detection and analysis of the *F. tularensis* 5', middle, and 3' regions of IS*Ftu1* transcript(s), as well as the analysis of genes surrounding three highly expressed IS*Ftu1* locations were performed according to the manufacturer's protocol for the AccessQuick RT-PCR System (Promega catalog number 9PIA170). The RT-PCR samples were prepared with 100 ng of RNA and amplified in a Veriti Thermocycler using a 45 min reverse transcription step at 45^{o}C followed immediately by a 2 min denaturing step at 95^{o}C. Finally, the cDNA was amplified using 35 cycles of 95^{o}C for 30 sec, 57^{o}C for 30 sec, and 72^{o}C for 30 sec. At the end of the cycling conditions, an extension step for 7 min at 72^{o}C was included. RT-PCR samples were stored at 4^{o}C for short term or -20^{o}C for long term.

First-strand cDNA synthesis of the co-transcribed genes encoding the universal stress protein and IS*Ftu1* in the *F. tularensis* type A.I strains during the various growth phases were prepared using the SuperScript III First-Strand Synthesis System, as recommended by the manufacturer (Invitrogen catalog number 18080-051). Briefly, 200 ng of *F. tularensis* RNA, dNTP mix containing 10 mM of each dNTP, the appropriate *F. tularensis* primer, and DEPC-treated water were combined, mixed, and incubated at 65^{o}C for 5 min, and then cooled on ice for 1 min. The cDNA Synthesis Mix was prepared according to manufacturer's recommendations and 10 µL of this mixture was added to each of the DNase-treated *F. tularensis* RNA samples. The reactions were then incubated at 50^{o}C for 50 min, followed by 85^{o}C for 5 min, and then chilled on ice. To remove the hybridized *F. tularensis* RNA from the cDNA, 1 µL of RNase H was added to each sample and mixed. The samples were then incubated at 37^{o}C for 20 min and stored at -20^{o}C or amplified immediately. Amplification of the cDNA was performed by using Platinum Taq, according to manufacturer's recommendations (Invitrogen catalog number 11304029).

The *F. tularensis* cDNA was denatured at 95ºC for 2 min and then amplified for 35 cycles with 95ºC for 30 sec, 51ºC for 30 sec, and 72ºC for 1 min. A final extension step of 72ºC for 7 mins was also performed. The *F. tularensis* amplified cDNA was kept at 4ºC for short term storage or -20ºC for long term storage.

Reactions for co-transcribed universal stress protein with IS*Ftu1* in *F. tularensis* Schu S4 were prepared according to manufacturer's protocol included with the OneStep RT-PCR Kit (Qiagen catalog number 210210). After preparing the reaction mixture and aliquoting this solution into each tube, 200 ng of the appropriate DNase-treated *F. tularensis* RNA was added. Using the Veriti Thermocycler, an initial reverse transcription step of 30 min at 50ºC followed immediately by a 15 min incubation at 95ºC step was performed. The initial step synthesized the first-strand cDNA, while the second step inactivated the reverse transcriptase, activated the HotStar Taq DNA polymerase, and denatured the cDNA. Lastly, the *F. tularensis* cDNA was amplified 35 cycles, using 94ºC for 30 sec, 51ºC for 30 sec, and 72ºC for 2 min. A final extension step of 72ºC for 10 min was also performed.

To check the amplified *F. tularensis* RT-PCR products, loading buffer was added to an aliquot of each sample and ran in a 1% agarose gel prepared with 0.5 g agarose in 50 mL 1x TAE buffer, and 1 µL of ethidium bromide, along with an appropriate dsDNA size marker. Agarose gels were ran at 90cV for approximately 1 h. Photographs of the ethidium bromide containing agarose gels were taken using the UV setting on the AlphaImager from ProteinSimple. The amplified *F. tularensis* RT-PCR reactions were kept at 4ºC for short term storage or -20ºC for long term storage.

5’ Rapid Amplification of cDNA Ends (RACE)

To determine the nucleotide sequence of the 5’ untranslated region for the transcript encoding the *F. tularensis* universal stress protein, the 5’/3’ RACE Kit, 2nd Generation (Roche catalog number 03 353 621 001) was used. These reactions were set up on ice and contained the FTT_0245-SP1 or IS*Ftu1*-SP1 primers and 2 µg of DNase-treated *F. tularensis* RNA. Reactions were incubated for 1 h at 55ºC to synthesize first-strand cDNA followed by 5 min at 85ºC. Next, 100 µL of Binding Buffer was added to each reaction. Samples were mixed and then transferred to a High Pure Filter Tube, which was provided in the kit, followed by centrifugation for 30 sec at 8,000 x g. The flow-through liquid was discarded and the filter containing the *F. tularensis* cDNA was washed by adding 500 µL of Wash Buffer and then centrifugation. The flow-through liquid was discarded and the filter containing the cDNA was washed again, using 200 µL of Wash Buffer this time followed by centrifugation for 2 min at 13,000 x g. The filter tube containing the purified first-strand cDNA was then transferred to a clean tube and eluted by adding 50 µL of Elution Buffer and centrifugation for 30 secs at 8,000 x g.

Next, the *F. tularensis* first-strand cDNA was poly(A)-tailed, according to manufacturer’s protocol. For this process, the sample was incubated for 3 min at 94ºC, chilled on ice, and then 1 µL of Terminal Transferase was added to each reaction. Samples were incubated at 37ºC for 25 min followed by 10 min at 70ºC. The *F. tularensis* poly(A)-tailed cDNA was then prepared for PCR amplification, according to manufacturer’s recommendations using the FTT_0245-SP2 or IS*Ftu1*-SP2 primers. The dA-tailed cDNA was amplified in the Veriti Thermocycler by first denaturing the sample at 94ºC for 2 min followed by 10 amplification cycles with 94ºC for 15 sec, 55ºC for 30 sec and 72ºC for 40 sec. Another round of amplification was performed for 25 cycles with 94ºC for 15 sec, 55ºC for 30 sec and 72ºC for 40 sec, along with the addition of 20 sec to each 72ºC elongation

step. A final elongation step of 72°C for 7 min was also performed. The amplified *F. tularensis* cDNA products were visualized on a 1% agarose gel along with appropriate dsDNA size marker, quantitated using the Qubit or Nanodrop, and then sent to Eurofins Genomics along with the FTT_0245-SP2 primer for subsequent sequencing.

In silico Analysis

F. tularensis nucleotide and protein sequences of interest were obtained from GenBank, the National Institute of Health (NIH) genetic sequence database that includes an annotated collection of all publicly available sequences, and is part of the International Nucleotide Sequence Database Collaboration, which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA) (Benson et al., 2013). The Basic Local Alignment Search Tool (BLAST), which is provided by the National Center for Biotechnology Information (NCBI) within the National Library of Medicine (NLM), was used to identify regions of similarity between *F. tularensis* nucleotide or protein sequences. Translation of *F. tularensis* IS*Ftu1* ORFs was done using EMBOSS Sixpack (Li et al., 2015) through the European Molecular Biology Laboratory European Bioinformatics Institute (EMBL-EBI https://www.ebi.ac.uk/). Protein alignment of IS*Ftu1* was performed using Clustal Omega (Sievers and Higgins, 2014) via EMBL-EBI. Protein structure of IS*Ftu1* was predicted using SWISS-MODEL (Arnold et al., 2006). Inverted repeats in IS*Ftu1* were identified using EMBOSS 6.3.1 palindrome software, and direct repeats were found using EMBOSS equicktandem.

F. tularensis RNA-Seq transcriptome data was normalized and compiled by Ufuk Nalbantoglu, (Croucher and Thomson, 2010; Giannoukos et al., 2012). After data normalization, the expression level of each gene was determined and annotated to the Schu S4 locus tags for ease of comparison of the same gene between the different *F.*

tularensis strains. The expression of IS*Ftu1* and adjacent genes for each *F. tularensis* strain was determined by visualization of BAM files from RNA-Seq transcriptome data in Integrative Genomics Viewer (IGV) from the Broad Institute (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). The number of reads for each gene was counted and evaluated using the following criteria. One to ten reads was considered to be low expression, 10-100 reads was considered to be moderate expression, and >100 reads was considered to be high expression. The minimum free energy mRNA secondary structure of the universal stress protein without and with IS*Ftu1* was determined using Sfold 2.2 free-ware (http://sfold.wadsworth.org/cgi-bin/index.pl).

Table 2: Primers used in RT-PCR, probe synthesis, and 5' RACE.

Primer Name	Oligonucleotide Sequence (5'-3')	Usage
ISFtu1-F206	ATGAAGAAGGTATGACGGAGTTCG	PCR amplification of IS*Ftu1* and DIG-labeled probes (see Figures 7A, 9, and 12)
FTT0245-F209	TTAGCGCGACCATTGGGAGG	PCR amplification of co-transcribed genes (see Figure 12)
ISFtu1-F210	GTGTTTAACAAAATACTCTCGG	PCR amplification of co-transcribed genes (see Figure 12)
FTT0272-F211	CCTAGTAGCTATTAGACAAAGGAG	PCR amplification of co-transcribed genes (see Figure 12)
FTT0821-F212	AAGATTGGGATTGATACAC	PCR amplification of co-transcribed genes (see Figure 12)
ISFtu1-mid-F213	TAGATGAGATGGGGTGTGAC	PCR amplification of middle IS*Ftu1* region and DIG-labeled probes (see Figures 7B, and 9)
ISFtu1-3'-F214	TATCTACCTCCGTACTCTCCAG	PCR amplification of 3' IS*Ftu1* region and DIG-labeled probes (see Figures 7C, and 9)
FTT_0245-F223	GTAGGGAGATTCCTAATGGCG	PCR amplification of co-transcribed genes and IS*Ftu1* (see Figures 18 and 19)
ISFtu1-F224	AATGCCATCATATAGCCAAG	PCR amplification of full-length sense and antisense IS*Ftu1* (see Figures 8B, 8C, and 18)
ISFtu1-R201	TAAAGCTAGCGACTCTGCCAC	PCR amplification of IS*Ftu1* and DIG-labeled probes (see Figures 7A, 9, and 12)
ISFtu1-R203	AATCTTGGCTATATGATGGC	PCR amplification of co-transcribed genes (see Figure 12)
FTT0248-R204	GCTTGATACTTACCACTATCAATCAC	PCR amplification of co-transcribed genes (see Figure 12)
FTT0274-R205	GCGGTACAAAGTGCTATTAAGATG	PCR amplification of co-transcribed genes (see Figure 12)
FTT0823-R206	GATTAATGATTAAACTTAGCCTCG	PCR amplification of co-transcribed genes (see Figure 12)
ISFtu1-mid-R207	CCTCACCATAACTAGGCTCACC	PCR amplification of middle IS*Ftu1* region and DIG-labeled probes (see Figures 7B, and 9)
ISFtu1-3'-R208	CCGAGAGTATTTTGTTAAACAC	PCR amplification of 3' IS*Ftu1* region and DIG-labeled probes (see Figures 7C, and 9)

ISFtu1-R217	GTTAGCCTAAACCTTTTCAATAGG	PCR amplification of full-length sense and antisense IS*Ftu1* and co-transcribed universal stress protein (see Figures 8B, 8C, 18, and 19)
FTT_0249-R218	CTACTACAACTAGAGCAGCCAG	PCR amplification of co-transcribed genes and IS*Ftu1* (see Figure 18)
FTT_0245-SP1	CTCCCAATAGCTCCTCTCCTATGGC	5’ RACE amplification (see Figure 16)
FTT_0245-SP2	GCATATTCAACTATCTCTGCTGCCGG	5’ RACE amplification (see Figure 16)
ISFtu1-SP1	GGATTTAAATCTGGAGAGTACGGAGGTAG	5’ RACE amplification (see Figure 16)
ISFtu1-SP2	GCTAAAACCTCACCATAACTAGGCTCACC	5’ RACE amplification (see Figure 16)

RESULTS

F. tularensis Genomic Content

The different strains of *F. tularensis* examined in this study are very similar with regards to genome size and % G+C content, as well as protein and gene totals. See Table 3 for a summary of genomic content for these various *F. tularensis* strains, which was obtained from NCBI. The *F. tularensis* genome size is between 1.89 and 1.91 million bases (Mb). Both subtype A.I and A.II strain have 32.3% G+C content, type B has 32.2% G+C content, while subsp. *novicida* has 32.5% G+C content. All *F. tularensis* strains examined contain 10 rRNA genes, with the exception of wild type A.II strain WY-00W4114 (hereafter referred to as W4114), which only contains 9 rRNA genes. All *F. tularensis* strains encode 38 tRNA genes.

The gene counts differed among the *F. tularensis* strains examined. *F. tularensis* A.I strain Schu S4 had the least amount of genes with 1,834 and wild type A.I strain NE061598 had 1,925. The A.II strain WY96-3418 (hereafter referred to as WY96) and wild type A.II W4114 had very similar gene counts at 1,919 and 1,920, respectively. Attenuated type B LVS contained the most genes with 1,964. The subsp. *novicida* strain U112 had 1,844 genes. Type A.I Schu S4 had the least predicted proteins with 1,556, followed by attenuated type B LVS with 1,580 predicted proteins. The wild type A.I NE061598 contained 1,662 predicted proteins and the A.II strains WY96 and W4114 have 1,647 and 1,643 predicted proteins, respectively. Subspecies *novicida* U112 contained the most predicted proteins with 1,743. The largest difference between the genomes in the *F. tularensis* strains was with the number of pseudogenes. *F. tularensis* subsp. *novicida* only contained 49 pseudogenes, while the *F. tularensis* subtype A.I, subtype A.II, and type B strains contained over 200 pseudogenes.

Table 3: *F. tularensis* strains used in this study.

F. tularensis	**RefSeq***	**Size (Mb)***	**% GC***	**Protein***	**rRNA***	**tRNA***	**Other RNA***	**Gene***	**Pseudo gene***
subsp. *tularensis* Schu S4	NC_006570.2	1.89	32.3	1,556	10	38	1	1,834	227
subsp. *tularensis* NE061598	NC_017453.1	1.89	32.3	1,662	10	38	4	1,925	221
subsp. *tularensis* WY96-3418	NC_009257.1	1.90	32.3	1,647	10	38	4	1,919	220
subsp. *tularensis* WY-00W4114	NZ_CP009753.1	1.90	32.3	1,643	9	38	4	1,920	226
subsp. *holarctica* LVS	NC_007880.1	1.90	32.2	1,580	10	38	4	1,964	322
subsp. *novicida* U112	NC_008601.1	1.91	32.5	1,743	10	38	4	1,844	49

*Data was obtained from NCBI (https://www.ncbi.nlm.nih.gov/genome/).

F. tularensis Growth Curves

Growth curves for five different *F. tularensis* strains were examined, including type A.I strain Schu S4, type A.I strain NE061598, type A.II strain WY96, type A.II strain W4114, attenuated type B (subsp. *holarctica*) LVS, and subsp. *novicida* strain U112. *F. tularensis* LVS and U112 reached exponential phase the quickest, doubling roughly every 1.4 hours, while WY96 had the longest lag phase, taking approximately 25 hours to begin early exponential growth phase and about 50 hours to reach stationary phase after inoculation (Figure 2). Schu S4, NE061598, and W4114 entered early exponential growth phase starting around 12 hours and reached stationary phase beginning after 30 or 35 hours (Figure 2). The virulent strains (Schu S4, NE061598, WY96, and W4114) doubled every 3 to 4 hours, taking approximately twice as long as the attenuated LVS and avirulent U112 strains. LVS and U112 reached an OD_{600} of about 2.7, while the type A strains only reached an OD_{600} of about 1.5, prior to stationary growth phase. *F. tularensis* is known to grow slow and requires supplementation for growth in liquid culture (Mc Gann et al., 2010), as was observed during this study.

Virulent *F. tularensis* strains are resistant to reactive oxygen and reactive nitrogen species, but the mechanism that this pathogen utilizes to combat these antimicrobial molecules is unknown (Newstead et al., 2014). To examine *F. tularensis* gene expression in response to reactive nitric oxide, SNAP was added to a liquid culture of A.I strain NE061598 30 minutes before isolating the RNA, during early exponential growth phase. SNAP generates nitric oxide and has been shown to contain *F. tularensis* in the macrophage phagosome and substantially inhibit *mglA* transcription, but not decrease the viability of *F. tularensis* (Tancred et al., 2011). MglA is a global regulator in *F. tularensis* and affects many genes, including genes in the *Francisella* pathogenicity island (Broms et al., 2010). SNAP was used in this study to determine if there was any effect on the

transcriptional profile of wild type A.I NE061598, with particular emphasis on the expression of the most abundant IS element, specifically IS*Ftu1*, which is described later. As shown in Figure 3, the addition of SNAP did not have any obvious effect on the growth of *F. tularensis* NE061598 compared to untreated cultures.

Spermine has been shown to upregulate IS elements in *F. tularensis* LVS and Schu S4 (Carlson et al., 2009). This study also proposed that spermine, which is present in the cytosol of macrophages, provides a cue for activation of *F. tularensis* gene expression that reduces the proinflammatory cytokine response. To evaluate the effect of spermine on the transcriptional expression of IS*Ftu1* and adjacent genes in wild type A.I strain NE061598, spermine was added to the liquid culture 30 minutes before isolation of RNA, during early exponential growth phase of this strain. The addition of spermine did not alter the growth of *F. tularensis* NE061598 compared to untreated cultures (Figure 3).

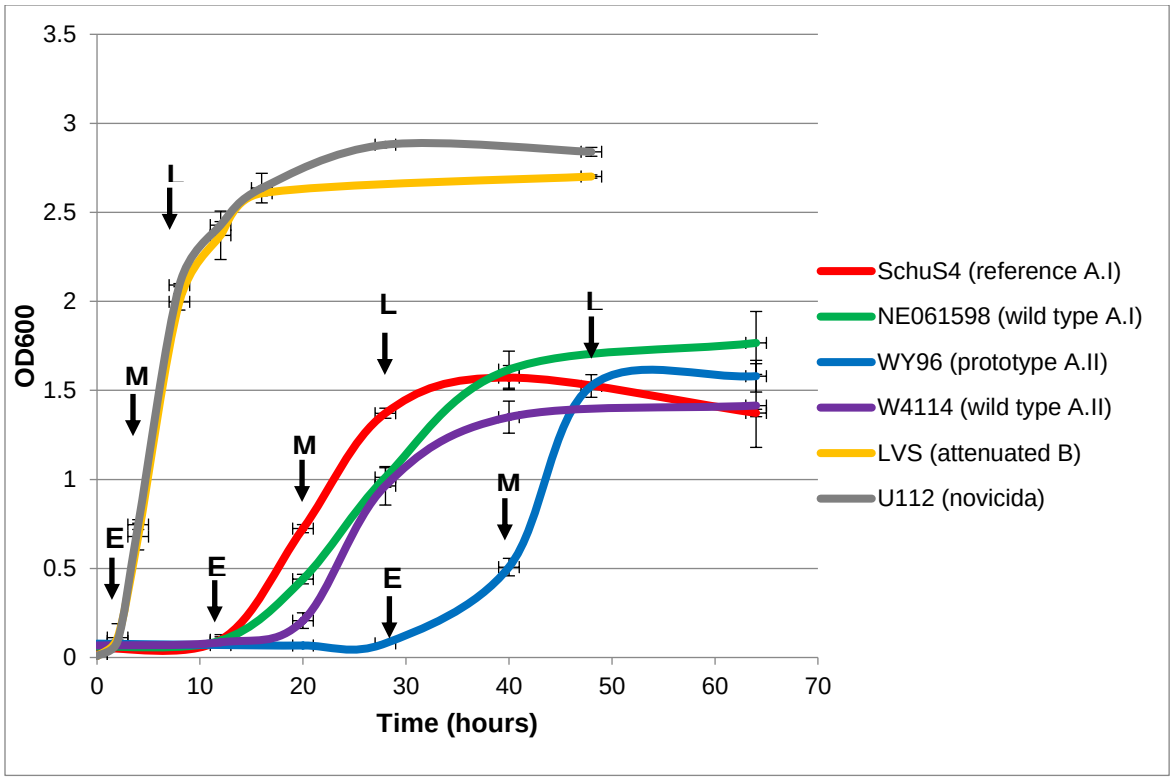


Figure 2: Growth curves of various *F. tularensis* strains. The average of absorbance readings at 600 nm from three independent experiments are shown throughout growth of the different *F. tularensis* strains, specifically Schu S4 (reference type A.I), NE061598 (wild type A.I), WY96 (prototype A.II), W4114 (wild type A.II), LVS (attenuated type B), and U112 (prototype subsp. *novicida*). All strains were grown in BHI supplemented with cysteine, β-NAD, Heme-Histidine and 20% glucose at 37ºC with CO_2 and with shaking at 200 rpm. Early (E), mid (M), and late (L) log growth phase are indicated with arrows, and standard deviations are denoted at the associated time points.

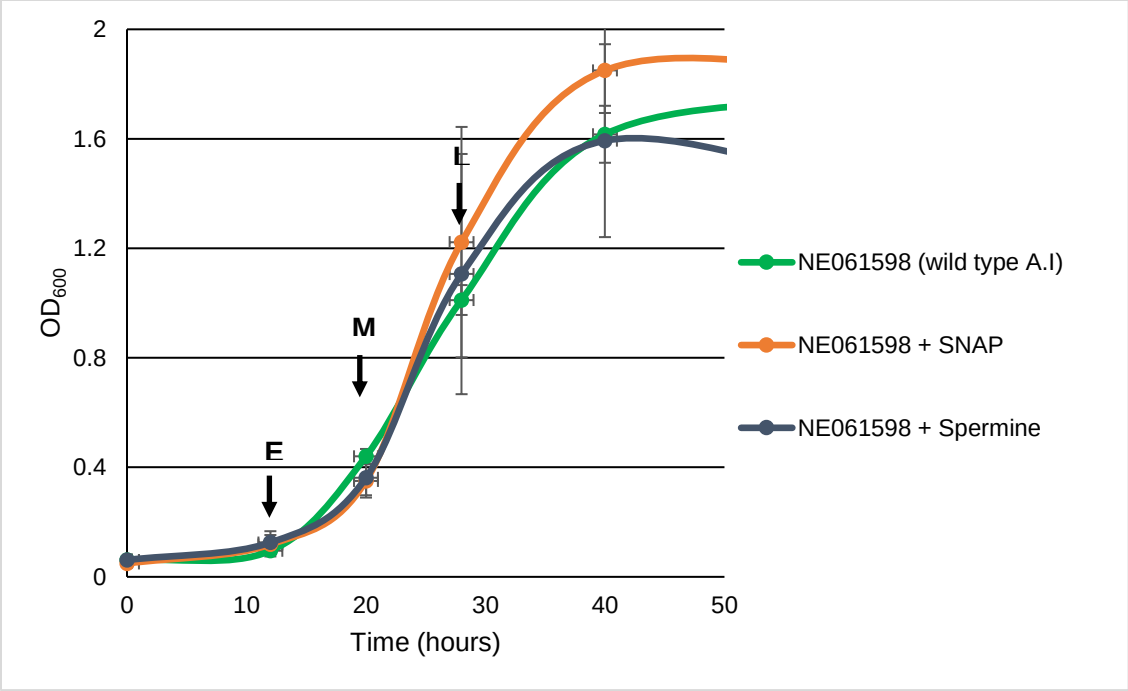


Figure 3: Growth of wild type A.I *F. tularensis* NE061598 in the absence or presence of nitric oxide producing SNAP or the polyamine spermine. The average of absorbance readings at 600 nm from three independent experiments are shown throughout growth of wild type A.I NE061598 grown without and with SNAP or spermine in BHI supplemented with cysteine, β-NAD, Heme-Histidine and 20% glucose at 37ºC with CO_2 and with shaking at 200 rpm. Early (E), mid (M), and late (L) log growth phase are indicated with arrows, and standard deviations are denoted at the associated time points.

IS*Ftu1* Open Reading Frame Analysis

The most abundant IS element in the virulent *F. tularensis* subspecies is IS*Ftu1* with 47 to 48 full-length copies in subspecies *tularensis* (type A), 51 copies in subspecies *holarctica* (type B), and 56 copies in subspecies *mediasiatica* (Table 1). In contrast, subspecies *novicida* has only one intact copy of IS*Ftu1*. The IS*Ftu1* open reading frame in type A.II and type B strains is 881 bp in size, whereas the type A.I strains contain a single nucleotide polymorphism, which produces a premature stop codon and a 785 bp open reading frame. The open reading frame for all IS*Ftu1* locations in *F. tularensis* subsp. *mediasiatica* is only 751 bp, due to a 130 bp deletion in the middle of the gene, after the -1 programmed translational frameshift. The deleted region does include the first aspartic acid in the catalytic DDE motif, indicating that an inactive transposase would be produced. *F. tularensis* subsp. *novicida* contains a single copy IS*Ftu1* that is 869 bp in size with a 12 bp deletion at the beginning of the gene, however the start codon is unaffected by this deletion.

The catalytic DDE motif that is present in the *F. tularensis* subtype A.I 785 bp, subtype A.II, and type B 881 bp IS*Ftu1* open reading frames, along with the -1 programmed translational frameshift slippery heptamer, are shown in Figure 4. The translated protein sequence from representative copies of IS*Ftu1* in *F. tularensis* Schu S4, WY96, and LVS were aligned to compare content using Clustal Omega (Figure 5A). Each amino acid difference is marked along with the DDE motif and premature stop codon in Schu S4 IS*Ftu1*. Compared to the IS*Ftu1* gene in WY96 and LVS, there were two non-conservative mutations in Schu S4 IS*Ftu1*, specifically a tyrosine to a histidine and a proline to a leucine, and one non-sense mutation in which tryptophan was mutated to a stop codon. There was only one non-conservative mutation in WY96 IS*Ftu1*, an aspartic acid to a tyrosine. In LVS IS*Ftu1*, there were two non-conservative mutations, an alanine

to a threonine and an arginine to an isoleucine, and two conservative mutations, a phenylalanine to a leucine and a lysine to an arginine. Figure 5B shows the predicted protein structure for IS*Ftu1* in LVS. A similar protein structure was obtained for *F. tularensis* A.I Schu S4 and A.II WY96 IS*Ftu1*; however, the Schu S4 IS*Ftu1* was slightly shorter due to the premature stop codon.

F. tularensis IS*Ftu1* is also surrounded by an inverted repeat and contains a direct repeat in the 5' untranslated region (UTR) (Figure 6). This inverted repeat and direct repeat is found in every *F. tularensis* subspecies, with the exception of *novicida* in which the IS*Ftu1* repeat is two nucleotides shorter and this gene does not have the direct repeat. However, in the type A.I strains, the direct repeat is repeated four times instead of two times. The IS*Ftu1* ORF in the type A genomes contains a 312 bp deletion in a few copies that is surrounded by a direct repeat, as shown in Figure 6. The 312 bp deletion appears in the genome three times for the type A.I strains Schu S4 and NE061598, and four times for the type A.II strains WY96 and W4114. This 312 bp deletion is not present in the IS*Ftu1* gene within the chromosome of attenuated type B LVS, subsp. *mediasiatica*, or subsp. *novicida*. A putative Rho independent transcriptional termination site is also present in *F. tularensis* IS*Ftu1* and is shown in Figure 6.

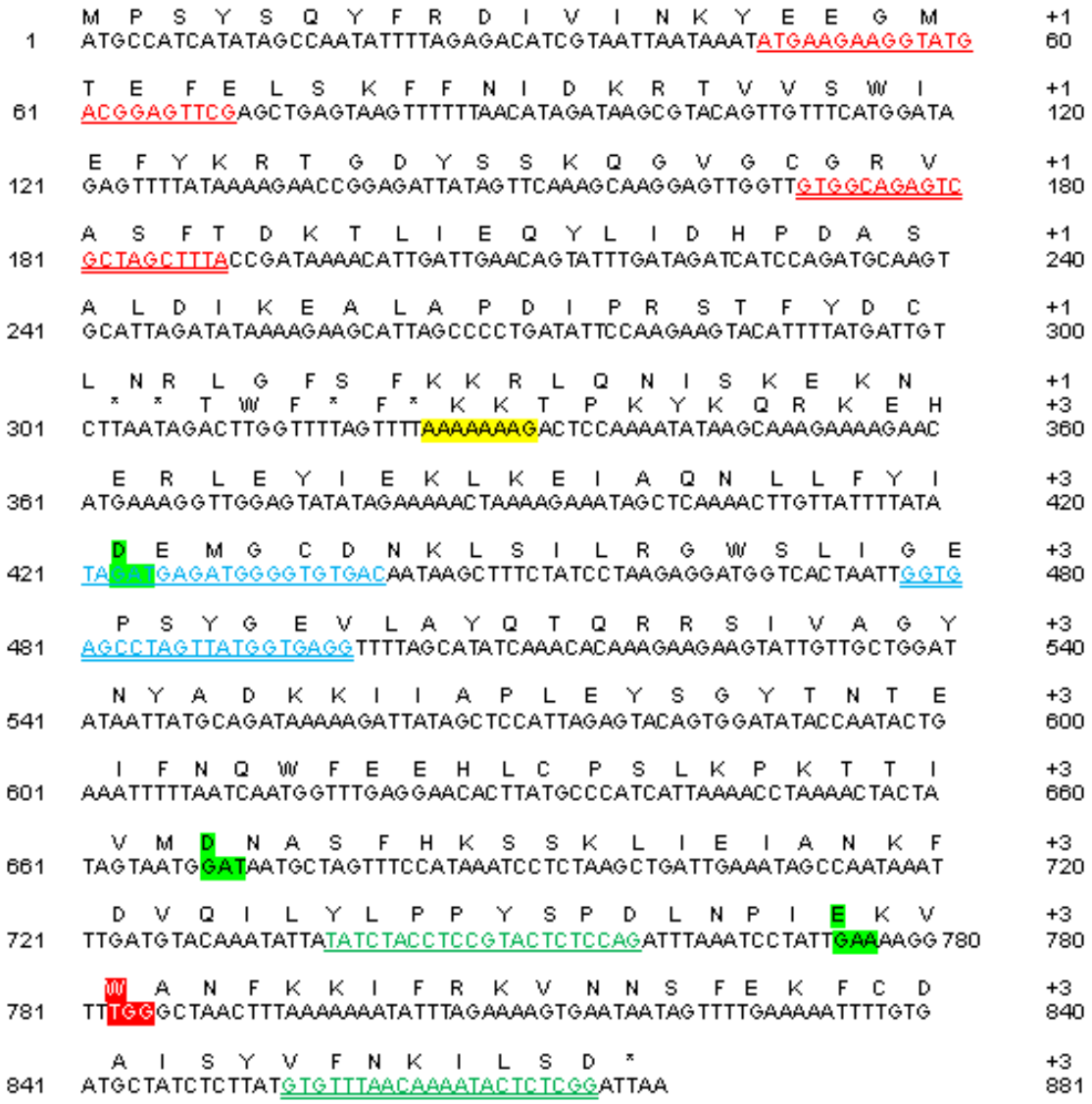


Figure 4: *F. tularensis* IS*Ftu1* open read frame, including nucleotide and protein sequence. Forward primers (underlined) and reverse primers (double underlined) are indicated that were used to initially detect the 5' region (red oligonucleotides), middle region (blue oligonucleotides), and 3' region (green oligonucleotides) of expressed IS*Ftu1* RNA by RT-PCR. The slippery nucleotide heptamer in the -1 programmed translational frameshift region is highlighted in yellow. The catalytic DDE motif is highlighted in green. The tryptophan codon, which is highlighted in red, is a stop codon in *F. tularensis* type A.I strains.

A

```
SchuS4   MPSYSQDFRDIVINKHEEGMTEFELSKFFNIDKRTVVSWIEFYKRTGDYS
WY96     MPSYSQYFRDIVINKYEEGMTEFELSKFFNIDKRTVVSWIEFYKRTGDYS
LVS      MPSYSQDFRDIVINKYEEGMTEFELSKFFNIDKRTVVSWIELYKRTGDYS

SchuS4   SKQGVGCGRVASFTDKTLIEQYLIDHPDASALDIKEALAPDIPRSTFYDCL
WY96     SKQGVGCGRVASFTDKTLIEQYLIDHPDASALDIKEALAPDIPRSTFYDCL
LVS      SRQGVGCGRVASFTDKTLIEQYLIDHPDASTLDIKEALAPDIPRSTFYDCL

SchuS4   NRLGFSFKKTPKYKQRKEHERLEYIEKLKEIAQNLLFYIDEMGCDNKLSIL
WY96     NRLGFSFKKTPKYKQRKEHERLEYIEKLKEIAQNLLFYIDEMGCDNKLSIL
LVS      NI LGFSFKKTPKYKQRKEHERLEYIEKLKEIAQNLLFYIDEMGCDNKLSIL

SchuS4   RGWSLIGEPSYGEVLAYQTQRRSIVAGYNYADKKIIALLEYSGYTNTEIF
WY96     RGWSLIGEPSYGEVLAYQTQRRSIVAGYNYADKKIIAPLEYSGYTNTEIF
LVS      RGWSLIGEPSYGEVLAYQTQRRSIVAGYNYADKKIIAPLEYSGYTNTEIF

SchuS4   NQWFEEHLCPSLKPKTTIVMDNASFHKSSKLIEIANKFDVQILYLPPYSPD
WY96     NQWFEEHLCPSLKPKTTIVMDNASFHKSSKLIEIANKFDVQILYLPPYSPD
LVS      NQWFEEHLCPSLKPKTTIVMDNASFHKSSKLIEIANKFDVQILYLPPYSPD

SchuS4   LNPIEKV * ANFKKIFRKVNNSFEKFCDAISYVFNKILSD*
WY96     LNPIEKV WANFKKIFRKVNNSFEKFCDAISYVFNKILSD*
LVS      LNPIEKV WANFKKIFRKVNNSFEKFCDAISYVFNKILSD*
```

B

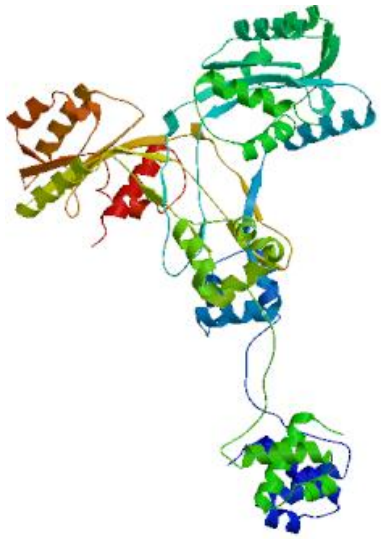


Figure 5: Alignment of representative open reading frames for the IS*Ftu1* protein in *F. tularensis* type A.I Schu S4, A.II strain WY96, and attenuated type B LVS. (A) The catalytic DDE motif is highlighted in green, the amino acid differences between each strain are highlighted in blue, and the premature stop codon in A.I strain Schu S4 is highlighted in red. ISFtu1 sequences were aligned using Clustal Omega (EMBL-EBI). **(B)** The predicted protein structure of ISFtu1 in F. tularensis LVS is shown.

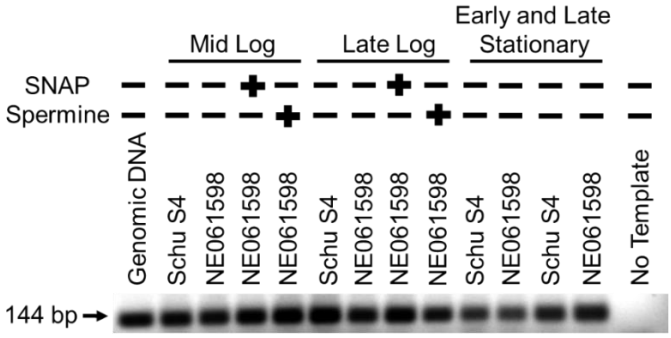
GTATATCGAAAATTACTAGATTTTATTCTAGATAGCTAGGCTATATTATGTAAAAAAATAGCCA
AAGCTTATATATAGTTAATCCGAAAGATATTTGTAGAAAAAGATATTTGTAGAAATGTTATAA
TGTCTAATAAAA**ATG**CCATCATATAGCCAATATTTTAGAGACATCGTAATTAATAAATATGAAG
AAGGTATGACGGAGTTCGAGCTGAGTAAGTTTTTTAACATAGATAAGCGTACAGTTGTTTCAT
GGATAGAGTTTTATAAAAGAACCGGAGATTATAGTTCAAAGCAAGGAGTTGGTTGTGGCAGA
GTCGCTAGCTTTACCGATAAAACATTGATTGAACAGTATTTGATAGATCATCCAGATGCAAGT
GCATTAGATATAAAAGAAGCATTAGCCCCTAATATTCCAAGAAGTACATTTTATGATTGTCTTA
ATAGACTTGGTTTTAGTTTTAAAAAAAGACTCCAAAATATAAGCAAAGAAAAGAACATGAAAG
GTTGGAGTATATAGAAAAACTAAAAGAAATAGCTCAAAACTTGTTATTTTATATAGATGAGAT
GGGGTGTGACAATAAGCTTTCTATCCTAAGAGGATGGTCACTAATTGGTGAGCCTAGTTATG
GTGAGGTTTTAGCATATCAAACACAAAGAAGAAGTATTGTTGCTGGATATAATTATGCAGATA
AAAAGATTATAGCTCCATTAGAGTACAGTGGATATACCAATACTGAAATTTTTAATCAATGGTT
TGAGGAACACTTATGCCCATCATTAAAACCTAAAACTACTATAGTAATGGATAATGCTAGTTT
CCATAAATCCTCTAAGCTGATTGAAATAGCCAATAAATTTGATGTACAAATATTATATCTACCT
CCGTATTCTCCAGATTTAAATCCTATTGAAAAGGTTTGGGCTAACTTTAAAAAAATATTTAGAA
AAGTGAATAATAGTTTTGAAAAATTTTGTGATGCTATCTCTTATGTGTTTAACAAAATACTCTC
GGATTAACTATAAACCAAGATAAATCAATATAGCTACTTTATAGGAAGTTAGCTTTAATTACTT
AAAAAGAATGGTAAATTTACATATAAATTATTCTATATACTAAGGTGTATGTTTTACATTTAAAT
TTCGGTAATATTAGATTTTAGTATAAAAAGGAAATAGTTGAATAATACCGCCTTTTTATACTC
ATTATTTTTT

Figure 6: Representative *F. tularensis* IS*Ftu1* gene showing direct and inverted repeats. Direct repeats in the 5' UTR of a representative *F. tularensis* IS*Ftu1* gene, from A.II strain WY96, are shown in green and blue text and are underscored with a dotted underline. The beginning and end of the IS*Ftu1* ORF is denoted with yellow highlighting. An inverted repeat flanking the IS*Ftu1* gene is shown in orange text with a wave underline, and an internal inverted repeat within the IS*Ftu1* gene is indicated with a thick underline. The slippery heptamer in the region with the -1 programmed translational frameshift is denoted in blue text. The gray highlighted 312 bp region that is deleted in some IS*Ftu1* open reading frames is flanked by direct repeats, which are in red text. An inverted repeat at the 3' region of the IS*Ftu1* gene is double-underlined and in green text, and may serve as a Rho independent transcriptional termination site.

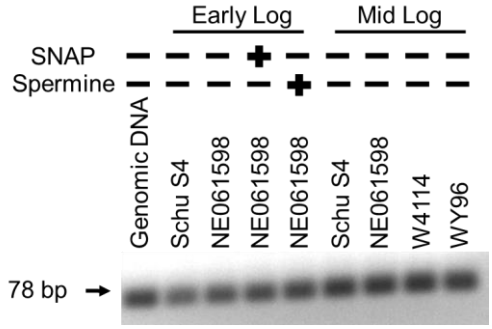
Reverse transcription PCR (RT-PCR) of IS*Ftu1* transcripts

Initial screening for *F. tularensis* IS*Ftu1* transcriptional expression was performed with high quality RNA and primers that would amplify the 5', middle, or 3' regions of this gene, using RT-PCR. See Table 2 for oligonucleotide sequences used and Figure 4 for the locations of these primers in IS*Ftu1*. Genomic DNA was used as the template for the positive controls with each primer pair and produced the expected 144 bp, 78 bp, and 139 bp PCR products for the 5', middle, and 3' regions of IS*Ftu1*, respectively. The negative control without template did not produce any amplicons as expected. RT-PCR with *F. tularensis* A.I strains Schu S4 and NE061598 cDNA produced the same amplicons as was obtained with the positive controls for the 5', middle, and 3' regions, during early, middle, and late log growth phase, as well as for early and late stationary phase (Figure 7A, 7B, and 7C). RT-PCR with cDNA derived from *F. tularensis* NE061598 exposed to SNAP or spermine also produced the same amplicons for early, middle, and late log growth phase. For the *F. tularensis* A.II strains WY96 and W4114, the expected bands for the middle and 3' regions during mid log phase were produced (Figure 7B and 7C). Although RT-PCR amplification of the cDNA from IS*Ftu1* produced the highest abundance of the 5' region and the lowest amount of the 3' region for this gene product, confirmation of these results using qRT-PCR is needed. Nevertheless, these results demonstrated that the A.I strains expressed the different regions of IS*Ftu1* throughout growth and during stationary phase, and that the A.II strains also transcribed these regions of IS*Ftu1* at least during mid log growth phase. These data also indicated that A.I NE061598 produced the IS*Ftu1* transcript after early log treatment with SNAP and spermine, similar to the results obtained without these treatments.

A



B



C

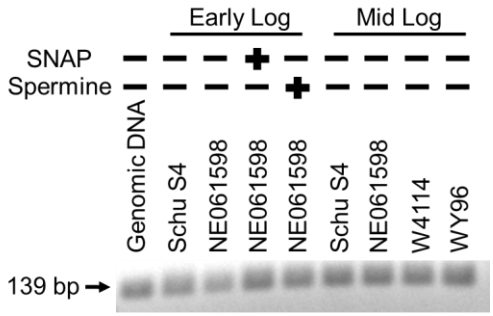


Figure 7: RT-PCR detection of 5', middle, and 3' region of expressed IS*Ftu1* RNA in *F. tularensis* A.I and A.II strains. Resulting RT-PCR products are shown following agarose gel fractionation and subsequent staining. The *F. tularensis* A.I strain, growth phase and if appropriate, treatment are shown above the associated lane. RT-PCR products obtained with primers that amplify **(A)** the 5' region of IS*Ftu1* cDNA derived from A.I strains Schu S4 and NE061598 at mid log, late log, and stationary growth phase without or with SNAP or spermine treatment, **(B)** the middle region of IS*Ftu1* cDNA derived from A.I strains Schu S4 and NE061598 at early log and mid log growth phase without or with SNAP or spermine treatment, along with A.II strains WY96 and W4114 at mid log growth phase, and **(C)** the 3' region of IS*Ftu1* cDNA derived from A.I strains Schu S4 and NE061598 at early log and mid log growth phase without or with SNAP or spermine treatment, along with A.II strains WY96 and W4114 at mid log growth phase. *F. tularensis* genomic. RT-PCR product lengths are denoted to the right of the respective gel photograph. Also shown are the RT-PCR results obtained for the positive control using *F. tularensis* genomic DNA as the template, as well as the negative controls that included a no template reaction and a no RTase reaction.

To determine if the *F. tularensis* full-length IS*Ftu1* transcript was expressed, the primers that flank the 5' and 3' region of this gene were used in RT-PCR with RNA isolated from the highly virulent type A.I strains Schu S4 and NE061598 (without or with SNAP or spermine added to the growth medium). The RNA isolated from *F. tularensis* Schu S4 and NE061598 and used in these analyses was confirmed to be of high quality, since the intensity of the 23S rRNA band was 2-fold higher than the 16S rRNA band, after agarose gel fractionation and ethidium bromide straining (Figure 8A).

Next, a gene-specific primer was used to generate first strand cDNA that is complementary to the sense strand of the IS*Ftu1* transcript in *F. tularensis* A.I strains Schu S4 and NE061598. RT-PCR was then performed with primers that amplify the full-length IS*Ftu1*. These results revealed that three RT-PCR products, which derived originally from the sense strand, were produced throughout growth and in stationary phase without or with SNAP or spermine treatment (Figure 8B). The expected full-length 880 bp amplicon was obtained for the RT-PCR samples using *F. tularensis* cDNA, as well as for the positive control that used genomic DNA as the template. However, a smaller 750 bp and 570 bp amplicon was also produced in the cDNA samples and genomic DNA positive control, but appeared to be considerably more abundant in the RT-PCR products (Figure 8B). The 570 bp cDNA derived amplicon corresponded to the IS*Ftu1* transcripts with the known 312 bp deletion that is in roughly 5% of the IS*Ftu1* genes within the *F. tularensis* genomes of type A strains. The faint 750 bp RT-PCR product was most likely due to non-specific amplification or possibly amplification of a RNA degradation product or truncated IS*Ftu1* transcript. The full-length 880 bp RT-PCR product and 570 bp RT-PCR product with the 312 bp deletion appeared to be similar in abundance for both Schu S4 and NE061598, throughout growth and during stationary phase. As expected, the no template and no RTase controls did not produce any amplification products (Figure 8B). These results

demonstrate that the full-length IS*Ftu1* transcript is being produced throughout the growth of the A.I strains Schu S4 and NE061598, along with the truncated 570 bp IS*Ftu1* transcript.

To investigate the directionality of the IS*Ftu1* transcript, a gene-specific primer that was complementary to antisense IS*Ftu1* transcripts was also used to generate first strand cDNA, since antisense is known to regulate IS elements and other genes that may have a deleterious effect on genome stability if produced (Thomason and Storz, 2010; Wagner et al., 2002). RT-PCR with this primer and another appropriate primer to amplify full-length antisense IS*Ftu1* transcripts in *F. tularensis* A.I strains Schu S4 and NE061598 again resulted in three RT-PCR products that were 880 bp, 750 bp, and 570 bp in length (Figure 8C). These RT-PCR products were produced throughout growth and in stationary phase without or with SNAP or spermine treatment (Figure 8C). Interestingly, there was considerably higher levels of the full-length antisense IS*Ftu1* transcript than the truncated 570 bp antisense IS*Ftu1* transcript (Figure 8C); however, the abundance of the full-length antisense IS*Ftu1* RT-PCR product appeared to be lower than the levels of the full-length sense IS*Ftu1* RT-PCR product (Figure 8B and 8C). Again the positive control containing *F. tularensis* genomic DNA and the no template and no RTase negative controls gave the expected results in the RT-PCR experiment to assess antisense IS*Ftu1*, as was obtained for the RT-PCR sense IS*Ftu1* results (Figure 8B and 8C). Together these results demonstrate that antisense transcript of IS*Ftu1* is being produced throughout the growth of both *F. tularensis* A.I strains Schu S4 and NE061598, albeit at apparently lower levels than the sense strand.

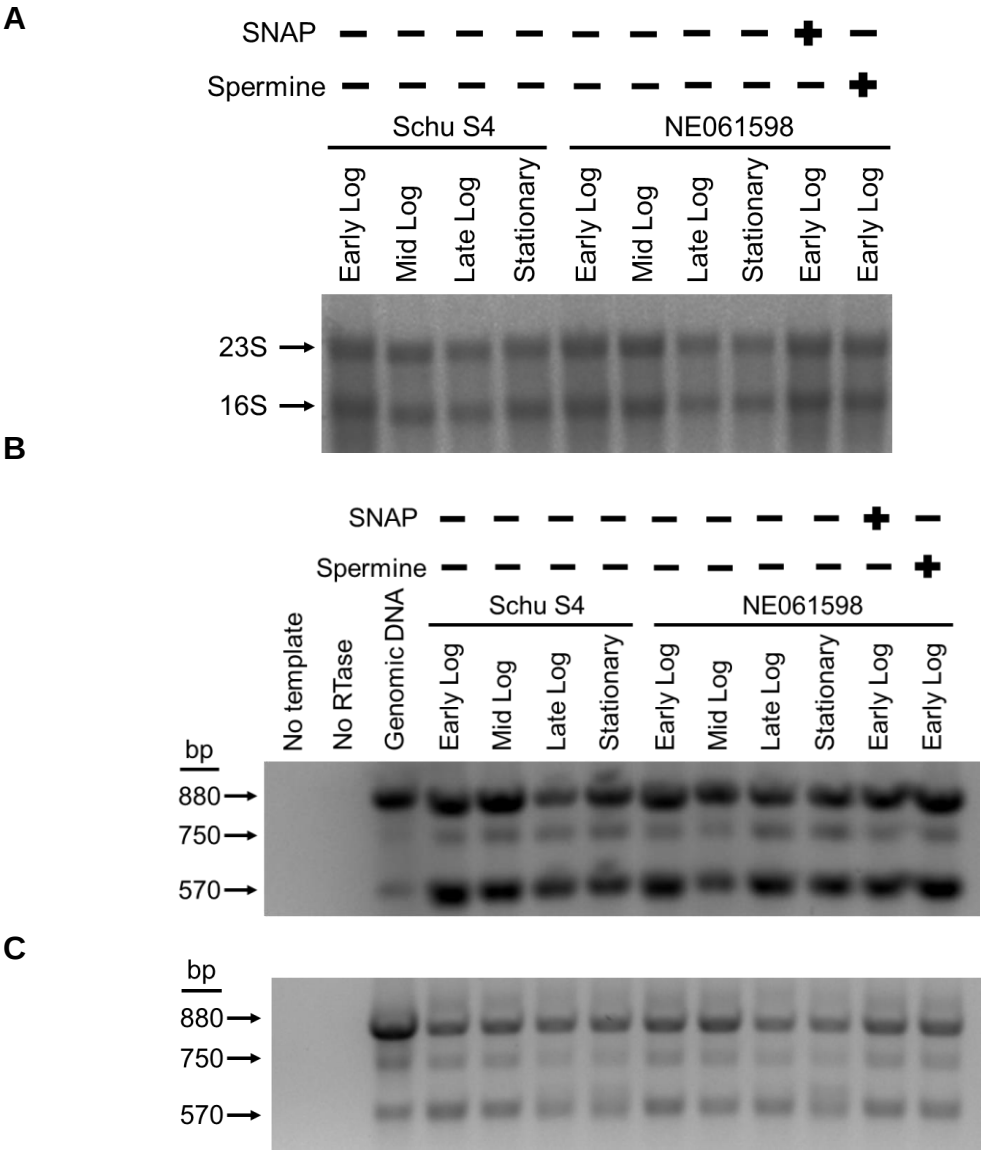


Figure 8: RT-PCR detection of full length and truncated IS*Ftu1* transcripts in *F. tularensis* type A.I strains Schu S4 and NE061598. (A) Shown is the fractionation of 1.5 µg per lane of total RNA isolated from the *F. tularensis* A.I strains throughout growth and stationary phase, without or with SNAP or spermine treatment, and subsequent staining. The 23S and 16S ribosomal RNA are denoted to the left. **(B and C)** Resulting RT-PCR products are shown following agarose gel fractionation and subsequent staining. The *F. tularensis* A.I strain, growth phase and if appropriate, SNAP or spermine treatment are shown above the associated lane. Also shown are the RT-PCR results obtained for the positive control using *F. tularensis* genomic DNA as the template, as well as the negative controls that included a no template reaction and a no RTase reaction. RT-PCR products obtain when first strand cDNA was synthesized using a primer complementary to the **(B)** sense strand or the **(C)** antisense strand of *F. tularensis* IS*Ftu1* RNA and then subsequent RT-PCR with primers that amplify the full-length gene product are shown.

Northern Blot Analysis of IS*Ftu1*

Northern blot analysis was performed to confirm the length and relative abundance of IS*Ftu1* transcripts in *F. tularensis* A.I strains Schu S4 and NE061598. For these assessments, probes were made to the 5', middle, and 3' region of IS*Ftu1* gene, using primers that amplify each of these regions and DIG labeling. Table 2 shows the oligonucleotide sequences and Figure 4 denotes the location in these primers in the IS*Ftu1* gene. Figure 9 shows the DIG-labeled IS*Ftu1* specific probes to the 3', middle, and 5' region of this gene, as well as the associated unlabeled 144 bp, 78 bp, and 139 bp amplicons, respectively. As expected, the DIG-labeled amplicons migrated slower than the unlabeled amplicons, due to the larger size and incorporation of DIG-11-dUTP versus dTTP into the labeled amplicon.

Prior to northern blot analysis, the integrity of the total RNA isolated from the *F. tularensis* A.I strains Schu S4 and NE061598 was checked by fractionation in an agarose gel and subsequent ethidium bromide staining. Figure 10A shows a representative photograph of the fractionated and ethidium bromide stained total RNA isolated from mid log *F. tularensis* A.I strains Schu S4 and NE061598, indicating that the RNA was of high quality. The IS*Ftu1* transcripts detected in northern blot analysis using the DIG-labeled probe to the 5' region of this gene were approximately 800 bp, 640 bp, 500 bp, and 320 bp in length from Schu S4 (Figure 10B, lane 1). The same transcripts were detected in total RNA isolated from NE061598 with the exception of the highest 800 bp transcript, which was not observed (Figure 10B, lane 2). These northern blot results indicated that the 500 bp IS*Ftu1* truncated transcript appeared to be the most abundant in both *F. tularensis* A.I strains (Schu S4 and NE061598), using the DIG-labeled probe to the 5' region of this gene.

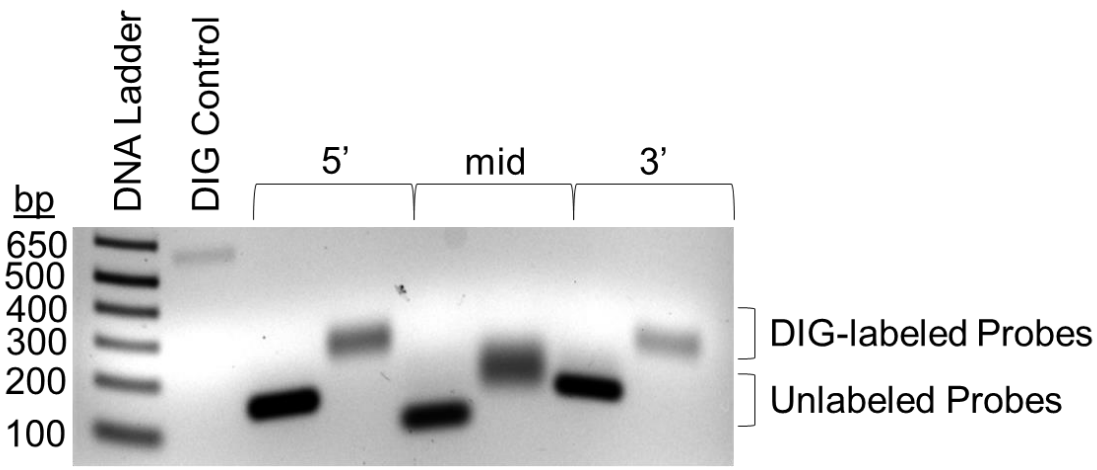
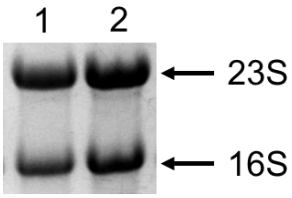


Figure 9: DIG-labeled *F. tularensis* IS*Ftu1* probes for northern blot analysis. Probes were made to the 5’, middle, and 3’ regions of IS*Ftu1,* using *F. tularensis* A.I strain NE061598 genomic DNA with the PCR DIG Probe Synthesis kit (Roche). The unlabeled probe and higher molecular weight DIG-labeled probe are shown for each IS*Ftu1* region, which is denoted above the respective lanes. A double-stranded DNA size marker and the DIG-labeled positive control, which was included in the kit, are also shown.

A



B

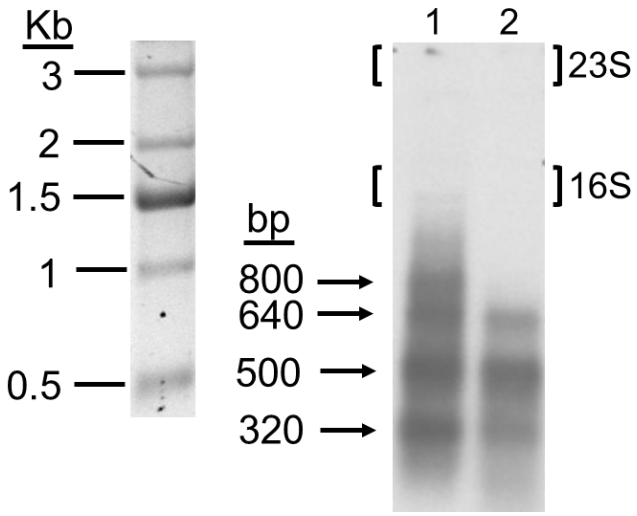


Figure 10: IS*Ftu1* RNA expression by *F. tularensis* A.I strains Schu S4 and NE061598. (A) Agarose gel fractionation and subsequent staining of total RNA that was isolated during mid exponential growth from *F. tularensis* type A.I strains Schu S4 (lane 1) and NE061598 (lane 2), prior to northern blot analysis. The location of the 23S and 16S ribosomal RNA are shown. **(B)** RNA size markers (left panel) and Northern blot analysis of IS*Ftu1* (right panel). using 5 µg of total RNA per lane that was isolated during mid exponential growth from *F. tularensis* type A.I strains Schu S4 (lane 1) and NE061598 (lane 2). DIG-labeled double-stranded DNA to the 5' region of IS*Ftu1* was used as the probe and the location of the 23S and 16S ribosomal RNA are denoted in brackets. The IS*Ftu1* transcripts observed were approximately 800 bp, 640 bp. 500 bp, and 320 bp in size. RNA size markers were used to estimate the length of these transcripts.

Northern blot analysis confirmed that IS*Ftu1* RNA transcripts are indeed being produced in both *F. tularensis* Schu S4 and NE061598. These results indicated that the 500 bp IS*Ftu1* transcript was expressed at a higher level than the full-length gene. However, the low sensitivity of northern blot analysis utilizing a non-radioactive probe did not allow for the detection of the full-length IS*Ftu1* transcripts in NE061598 that was observed in RT-PCR, but did detect an abundant 320 bp transcript that was not observed in RT-PCR. Importantly, northern blot analysis revealed numerous transcripts that differed in length contained at least the 5' end of IS*Ftu1*, since the probe that was utilized was specific to this region.

Differential Expression and Co-Expression of IS*Ftu1* with Adjacent Genes

To determine if genes adjacent to each IS*Ftu1* were co-expressed in *F. tularensis* during mid log growth phase, deep RNA sequencing was performed using the Illumina RNA-Seq technology and the associated BAM files were assessed. Each IS*Ftu1* along with the upstream and downstream adjacent genes were examined for co-transcription with IS*Ftu1* in *F. tularensis* subtype A.I strains Schu S4 and NE061598 (without and with SNAP or spermine), subtype A.II strains WY96 and W4114, type B LVS, and subsp. novicida U112. RNA expression levels were determined by the number of reads mapped to each gene, and the relative level of transcript abundance was designated as high, moderate, low or no expression, as shown in Tables 4-11. More specifically, 10 or less reads was considered low expression, 10-100 reads corresponded to moderate expression, and everything over 100 reads was recorded as high expression. Due to the identical or similar nucleotide content of each IS*Ftu1* ORF, RNA expression was analyzed for the 5' and 3' regions of each IS*Ftu1*, since those reads contained unique sequences that continued beyond the IS*Ftu1* gene and allowed for mapping of their location in the *F. tularensis* chromosome.

As summarized in Table 12, the majority of 5' IS*Ftu1* expression in *F. tularensis* Schu S4, NE061598, WY96, and LVS was low, whereas moderate expression of this region was predominately observed in W4114. The single IS*Ftu1* ORF in U112 had moderate expression, but there was no co-expression of the 5' or 3' adjacent gene (Table 9).

The addition of SNAP to NE061598 during growth did increase the expression of IS*Ftu1* and the adjacent genes, as shown in Table 10 and summarized in Table 12. There was higher IS*Ftu1* expression in both the 5' and 3' ends, and there were also more 5' and 3' adjacent genes with low, moderate, and high expression compared to NE061598 without SNAP. Likewise, the addition of spermine to NE061598 also increased the expression of IS*Ftu1* in the 5' and 3' regions of this gene, as well as the 5' and 3' adjacent genes. A comparison of the relative number of IS*Ftu1* and adjacent genes that were upregulated in response to SNAP and spermine showed that both SNAP and spermine stimulated higher expression levels overall. Collectively, these findings further validated that IS*Ftu1* is expressed and that the expression of this gene is responsive to external factors such as nitric oxide produced by SNAP and the polyamine spermine. These results also suggest that transcription of IS*Ftu1* may affect the expression of adjacent genes.

Table 4: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* A.I Schu S4 during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the Schu S4 genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product (gene)**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product (gene)**
FTT_0005 (4086-4625)	succinate-semialdehyde dehydrogenase, fragment (*gabD1*)	x	+	FTT_0004c (3204-3988)	+	+	FTT_0003c (2692-3096)	conserved hypothetical membrane protein, fragment
FTT_0008 (7959-8252)	sugar transferase, fragment	++	+	FTT_0009 (8350-8822)	+	+	FTT_0010 (8924-9430)	modification methylase, fragment
FTT_0099 (103101-103844)	transposase (IS*Ftu2*)	x	+	FTT_0098c (102215-102999)	+	+	FTT_0097 (101596-102141)	hypothetical protein
FTT_0132 (144667-146199)	anaerobic glycerol-3-phosphate dehydrogenase (*glpA*)	++	++	FTT_0131c (143567-144351)	+	++	FTT_0130 (141966-143474)	glycerol kinase (*glpK*)
FTT_0170c (185497-185925)	conserved hypothetical protein	+	+	FTT_0171 (185991-186775)	x	+	FTT_0172 (186876-187226)	hypothetical membrane protein, fragment
FTT_0215 (231974-234127)	primosomal protein N' (*priA*)	+	+	FTT_0216 (234257-235041)	x	+	FTT_0217 (235142-235378)	hypothetical protein, pseudogene
FTT_0227c (243546-243740)	hypothetical protein	++	+	FTT_0226c (242738-243522)	+	x	FTT_0225c (241608-242639)	MFS transport protein, fragment
FTT_0246c (259287-259502)	hypothetical protein, pseudogene	+++	+++	FTT_0247 (259568-260352)	+	x	FTT_0248 (260522-260893)	hypothetical protein

FTT_0254c (270392-270763)	hypothetical protein	x	+	FTT_0253c (269465-270249)	x	+++	FTT_0252 (267995-269368)	2-isopropylmalate synthase (*leuA*)
FTT_0272 (287222-287542)	hypothetical membrane protein	+++	++	FTT_0273 (287566-288350)	+	+	FTT_0274 (288464-288682)	hypothetical membrane protein
FTT_0351 (350614-351051)	50s ribosomal protein L7 (*rplQ*)	+++	+	FTT_0352 (351205-351989)	+	x	FTT_0353c (352087-352548)	transposase, fragment (IS*Ftu2*)
FTT_0356 (355151-357037)	chaperone Hsp90, heat shock protein HtpG (*htpG*)	++	+	FTT_0355c (354205-354989)	x	+++	FTT_0354 (352918-353913)	hypothetical protein
FTT_0358 (358209-359365)	conserved hypothetical protein, pseudogene	x	+	FTT_0357c (357362-358145)	x	+++	FTT_0356 (355151-357037)	chaperone Hsp90, heat shock protein HtpG (*htpG*)
FTT_0362c (362474-363346)	hypothetical protein	x	+	FTT_0363 (363514-364298)	+	++	FTT_0364c (364422-364871)	hypothetical protein
FTT_0376c (377481-378539)	hypothetical membrane protein	x	+	FTT_0377 (378906-379690)	+	+++	FTT_0378c (379982-380977)	hypothetical protein
FTT_0441c (454334-454974)	NADH dehydrogenase subunit, pseudogene	x	x	FTT_0440c (453349-454133)	+	++	FTT_0439 (452619-453275)	tRNA/rRNA methyltransferase (*yjfH*)
FTT_0514 (535309-536435)	L-lactate dehyrogenase, pseudogene (*lldD2*)	+	x	FTT_0513c (534457-535241)	+	+++	FTT_0512 (533805-534353)	SNO glutamine amidotransferase family protein
FTT_0560c (577004-578056)	phosphoserine aminotransferase (*serC*)	+	+	FTT_0561 (578376-579160)	+	++	FTT_0562 (579542-580663)	polyamine transporter, ABC transporter, ATP-binding protein (*potG*)
FTT_0693c (711592-712416)	formamidopyrimidine-DNA glycosylase (*mutM*)	+	+	FTT_0692c (710737-711521)	+	++	FTT_0691 (709046-710236)	tyrosyl-tRNA synthetase (*tyrS*)
FTT_0723c (745380-746168)	short-chain dehydrogenase/reductase family protein	x	+	FTT_0722c (744289-745073)	x	+++	FTT_0721c (741200-743425)	peroxidase/catalase (*katG*)

FTT_0764 (782778-784862)	glycyl-tRNA synthetase subunit beta (*glyS*)	++	+	FTT_0765 (785002-785786)	+	+++	FTT_0766 (785955-786677)	purine nucleoside phosphorylase (*deoD*)
FTT_0778 (798783-799157)	hypothetical protein	+	+	FTT_0779 (799235-800019)	x	+	FTT_0780c (800295-800564)	hypothetical membrane protein
FTT_0821 (838221-838925)	hypothetical protein	+	++	FTT_0822 (839217-840001)	+	++	FTT_0823 (840225-840872)	MutT protein
FTT_0919 (929356-930801)	hypothetical protein	++	+	FTT_0920 (930936-931720)	+	+	FTT_0921 (932217-932687)	pseudogene
FTT_1066c (1074364-1074738)	hypothetical protein	x	+	FTT_1065c (1073411-1074195)	+	++	FTT_1064 (1072282-1072872)	pseudogene
FTT_1073c (1084830-1085343)	pseudogene	x	+	FTT_1074 (1085435-1086219)	+	+	FTT_1075 (1087103-1087330)	transcriptional regulator
FTT_1143 (1155021-1155266)	hypothetical protein	x	+	FTT_1142c (1153900-1154684)	+	+++	FTT_1141 (1153701-1154956)	pseudogene
FTT_1211c (1230784-1231176)	hypothetical protein	++	++	FTT_1210c (1229938-1230722)	+	+	FTT_1209c (1227915-1229837)	pseudogene
FTT_1231 (1249168-1250364)	PP-loop family protein (*mesJ*)	++	+	FTT_1232 (1250509-1251293)	+	++	FTT_1233c (1251367-1252695)	proton-dependent oligopeptide transport (POT) family protein (*yjdL*)
FTT_1261c (1283043-1283614)	pseudogene	x	+	FTT_1262 (1283965-1284749)	+	+++	FTT_1263c (1285041-1286036)	hypothetical protein
FTT_1266c (1287773-1288486)	pirin (*yhhW*)	++	++	FTT_1265c (1286965-1287749)	x	+	FTT_1264 (1286404-1286891)	transposase (IS*Ftu2*)
FTT_1310c (1332931-1334847)	ATP-dependent metalloprotease (*hflB*)	+++	+++	FTT_1309c (1332221-1332693)	+	++	FTT_1308c (1331184-1331792)	hypothetical protein

FTT_1364 (1408560-1410035)	pseudogene (*treA*)	+	+	FTT_1363c (1407396-1408180)	+	x	FTT_1362 (1406420-1408582)	pseudogene
FTT_1385c (1429699-1429968)	hypothetical protein	+	+	FTT_1386 (1429959-1430743)	++	++	FTT_1387c (1431189-1433225)	DNA ligase (*ligN*)
FTT_1426c (1475548-1475847)	hypothetical protein	x	+	FTT_1427 (1476022-1476806)	+	+	FTT_1428c (1476967-1477482)	acetyltransferase
FTT_1450c (1499039-1500085)	dTDP-D-glucose 4,6-dehydratase (*wbtM*)	++	+	FTT_1449c (1498410-1498894)	+	++	FTT_1448c (1496666-1498072)	mannose-1-phosphate guanylyltransferase (*manC*)
FTT_1482 (1532644-1532862)	pseudogene	++	+	FTT_1481c (1531728-1532512)	+	x	FTT_1479c (1530815-1531003)	hypothetical protein
FTT_1547 (1609288-1609770)	pseudogene	x	+	FTT_1548 (1609794-1610578)	+	+	FTT_1549 (1611289-1611744)	hypothetical protein
FTT_1585 (1649614-1649869)	pseudogene	+	++	FTT_1584c (1648839-1649623)	x	+	FTT_1583 (1648373-1648765)	pseudogene
FTT_1600c (1666398-1667912)	fumerate hydratase (*fumA*)	+++	+	FTT_1599c (1665129-1665913)	+	+	FTT_1598 (1662198-1665032)	hypothetical protein
FTT_1613 (1676756-1676992)	BolA-like protein	+	++	FTT_1615 (1678189-1679973)	++	++	FTT_1616 (1679445-1680839)	cysteinyl-tRNA synthetase (*cysS*)
FTT_1619 (1682801-1683289)	pseudogene	x	++	FTT_1620 (1683438-1684222)	+	++	FTT_1621c (1684319-1685482)	pseudogene
FTT_1642c (1706937-1707626)	pseudogene (*hsdR3*)	x	+	FTT_1643 (1707692-1708476)	+	x	FTT_1644 (1708832-1710256)	Beta-fructofuranosidase
FTT_1668 (1736516-1737739)	serine transporter (*sdaC2*)	x	+	FTT_1669 (1738723-1739507)	+	+	FTT_1670c (1739584-1740195)	pseudogene (*napH*)

FTT_1691 (1760372-1760896)	hypothetical protein	+	++	FTT_1692 (1760920-1761704)	+	++	FTT_1693c (1761847-1762611)	hypothetical protein
FTT_1698c (1766954-1767670)	formate dehydrogenase (*fdh*)	x	++	FTT_1697c (1766112-1766896)	x	+++	FTT_1696 (1764262-1765896)	molecular chaperone GroEL (*groEL*)
FTT_1719c (1801580-1802095)	pseudogene	++	+	FTT_1718c (1800731-1801515)	x	x	FTT_1717 (1799755-1800634)	pseudogene
FTT_1743 (1832700-1833197)	pseudogene	x	+	FTT_1742c (1831850-1832634)	+	+	FTT_1741c (1831310-1831750)	pseudogene
FTT_1757c (1845684-1846939)	pseudogene	+	+	FTT_1756c (1844748-1845532)	x	+	FTT_1755 (1843755-1844671)	pseudogene (*gabD2*)
FTT_1779 (1868146-1868747)	pseudogene	x	+	FTT_1780 (1868806-1869590)	+	++	FTT_1781c (1869829-1870578)	hypothetical protein

Table 5: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* A.I NE061598 during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the NE061598 genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
NE061598_RS09860 (3957-4178)	hypothetical protein	x	+	NE061598_RS00020c (3204-3988)	x	+	NE061598_RS00015c (2692-3096)	hypothetical protein
NE061598_RS00040 (8039-8374)	hypothetical protein	+	+	NE061598_RS00045 (8446-8915)	x	x	NE061598_RS00055 (9017-9526)	restriction endonuclease
NE061598_RS00535 (103197-103940)	IS5 family transposase IS*Ftu2*	x	+	NE061598_RS09885c (102311-103095)	+	+	NE061598_RS00520 (101692-102237)	hypothetical protein
NE061598_RS09905 (144438-144536)	rRNA methyltransferase	++	+	NE061598_RS09900c (143663-144447)	x	++	NE061598_RS00695 (142062-143570)	glycerol kinase
NE061598_RS00925c (185593-186057)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	x	+	NE061598_RS09910 (186087-186871)	+	+	NE061598_RS00940 (187050-187322)	hypothetical protein
NE061598_RS09920c (234207-234362)	hypothetical protein	x	+	NE061598_RS09925 (234353-235137)	+	+	NE061598_RS01180 (235205-235474)	hypothetical protein
NE061598_RS01240c (243642-243836)	hypothetical protein	++	+	NE061598_RS09930c (242834-243618)	x	x	NE061598_RS01225c (241704-242735)	MFS transporter
NE061598_RS09935c (259599-259673)	50S rRNA methyltransferase	+++	+	NE061598_RS09940 (259664-260448)	x	x	NE061598_RS01350 (260618-260989)	hypothetical protein
NE061598_RS01385c	hypothetical protein	+	+	NE061598_RS09950c	x	+++	NE061598_RS01370	2-isopropylmalate synthase

(270488-270859)				(269561-270345)			(268091-269464)	
NE061598_RS01480 (287318-287638)	hypothetical protein	x	x	NE061598_RS09960 (287662-288446)	x	x	NE061598_RS01500c (288520-289639)	MFS transporter
NE061598_RS01890 (350683-351120)	50S ribosomal protein L17	++	+	NE061598_RS09965 (351274-352058)	x	x	NE061598_RS01905c (352132-352726)	IS5/IS1182 family transposase
NE061598_RS01930 (355425-356483)	hypothetical protein	x	+	NE061598_RS09970c (354274-355058)	+	+++	NE061598_RS01910 (352987-353982)	hypothetical protein
NE061598_RS02005 (370618-371490)	hypothetical protein	x	+	NE061598_RS09985c (369666-370450)	+	++	NE061598_RS01990 (369093-369542)	nuclear transport factor 2 family protein
NE061598_RS09990c (375691-375828)	hypothetical protein	x	+	NE061598_RS09995 (375819-376603)	+	+++	NE061598_RS02040c (376927-378813)	molecular chaperone HtpG
NE061598_RS02040c (376927-378813)	molecular chaperone HtpG	+++	++	NE061598_RS10000 (378975-379759)	x	+++	NE061598_RS02055c (380051-381046)	hypothetical protein
NE061598_RS02375c (454399-454674)	hypothetical protein	x	+	NE061598_RS10010c (453417-454201)	+	++	NE061598_RS02360 (452687-453343)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
NE061598_RS02750 (535404-536503)	alpha-hydroxy-acid oxidizing enzyme	x	+	NE061598_RS10025c (534525-535309)	x	++	NE061598_RS02735 (533873-534421)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
NE061598_RS02995c (578188-578442)	hypothetical protein	x	+	NE061598_RS10045 (578433-579217)	x	x	NE061598_RS03010 (579599-580720)	polyamine ABC transporter ATP-binding protein
NE061598_RS03700c (711649-712473)	formamidopyrimidine-DNA glycosylase	+	+	NE061598_RS10090c (710794-711578)	x	+	NE061598_RS03685 (709103-710293)	tyrosine--tRNA ligase
NE061598_RS10100 (745121-745270)	hypothetical protein	x	+	NE061598_RS10095c (744346-745130)	x	+++	NE061598_RS03850c (741257-743482)	catalase-peroxidase
NE061598_RS04100		++	+	NE061598_RS10110	+	+++	NE061598_RS04115	

(782835-784919)	glycine--tRNA ligase subunit beta			(785059-785843)			(786012-786734)	purine-nucleoside phosphorylase
NE061598_RS10120c (799182-799301)	rRNA methyltransferase	x	+	NE061598_RS10125 (799292-800076)	x	x	NE061598_RS10130 (800121-800363)	hypothetical protein
NE061598_RS10140 (839071-839250)	peptidase M13	+	+	NE061598_RS10145 (839274-840058)	+	+	NE061598_RS04435 (840282-840929)	DNA mismatch repair protein MutT
NE061598_RS04930 (929421-930866)	DUF3573 domain-containing protein	+	+	NE061598_RS10150 (931001-931785)	+	x	NE061598_RS04945c (932093-932272)	hypothetical protein
NE061598_RS05735c (1074428-1074802)	hypothetical protein	x	x	NE061598_RS10170c (1073475-1073891)	x	x	NE061598_RS05720 (1072346-1072936)	histidine-type phosphatase
NE061598_RS05780c (1084894-1085265)	hypothetical protein	x	+	NE061598_RS10175 (1085499-1086283)	+	x	NE061598_RS05795c (1086693-1086902)	hypothetical protein
NE061598_RS10185 (1153765-1153890)	hypothetical protein	x	+	NE061598_RS10190c (1153964-1154748)	x	+++	NE061598_RS10195 (1154739-1155020)	hypothetical protein
NE061598_RS06510c (1230848-1231240)	hypothetical protein	++	+	NE061598_RS10210c (1230002-1230786)	x	+	NE061598_RS06495c (1227979-1229961)	peptidase M13
NE061598_RS06610 (1249232-1250428)	tRNA(Ile)-lysidine synthetase	+	+	NE061598_RS10220 (1250573-1251357)	x	++	NE061598_RS06625c (1251431-1252759)	MFS transporter
NE061598_RS10230c (1283681-1283875)	hypothetical protein	x	+	NE061598_RS10235 (1283869-1284653)	x	+++	NE061598_RS06790c (1284945-1285940)	hypothetical protein
NE061598_RS06810c (1287693-1288406)	pirin family protein	+	+	NE061598_RS10240c (1286869-1287653)	x	x	NE061598_RS06795 (1286201-1286795)	IS5/IS1182 family transposase
NE061598_RS07055 (1341129-1341386)	hypothetical protein	x	+	NE061598_RS10245c (1340197-1340981)	+	++	NE061598_RS07040 (1339291-1340100)	MFS transporter
NE061598_RS07165c	hypothetical protein	x	+	NE061598_RS10255	x	++	NE061598_RS10260c	hypothetical protein

(1362500-1362769)				(1362760-1363544)			(1363614-1363993)	
NE061598_RS07405c (1408369-1408668)	transcriptional regulator	x	+	NE061598_RS10275 (1408843-1409627)	x	x	NE061598_RS07420c (1409788-1410303)	N-acetyltransferase
NE061598_RS10295 (1465324-1465683)	hypothetical protein	+	+	NE061598_RS10290c (1464549-1465333)	x	x	NE061598_RS07700c (1464209-1464481)	phosphatase PAP2 family protein
NE061598_RS08060 (1542117-1542599)	cyclopropane-fatty-acyl-phospholipid synthase	x	+	NE061598_RS10310 (1542623-1543407)	x	x	NE061598_RS08075 (1543475-1544161)	class I SAM-dependent methyltransferase
NE061598_RS08285 (1582443-1582682)	hypothetical protein	x	+	NE061598_RS10315c (1581668-1582452)	x	x	NE061598_RS08265 (1581193-1581661)	IS1595 family transposase
NE061598_RS08360c (1599211-1600725)	fumarate hydratase	+	+	NE061598_RS10320c (1597942-1598726)	x	+	NE061598_RS08345 (1595011-1597845)	hypothetical protein
NE061598_RS08465 (1615614-1616105)	GNAT family N-acetyltransferase	x	+	NE061598_RS10335 (1616251-1617035)	x	++	NE061598_RS08480c (1617444-1618295)	hypothetical protein
NE061598_RS08590c (1639747-1640511)	type I restriction endonuclease subunit R	x	+	NE061598_RS10345 (1640505-1641289)	+	x	08605NE061598_RS (1641645-1643069)	beta-fructofuranosidase
NE061598_RS10355c (1671285-1671456)	endoglucanase	x	+	NE061598_RS10360 (1671536-1672320)	x	x	NE061598_RS08740c (1672394-1673008)	sodium:proton antiporter
NE061598_RS08850 (1693185-1693709)	hypothetical protein	x	+	NE061598_RS10365 (1693733-1694517)	x	+	NE061598_RS08865c (1694660-1695454)	peroxide stress protein YaaA
NE061598_RS08900c (1699767-1700483)	formate dehydrogenase	++	+	NE061598_RS10375c (1698925-1699709)	+	+++	NE061598_RS08885 (1697075-1698709)	molecular chaperone GroEL
NE061598_RS10400 (1801438-1801557)	hypothetical protein	++	+	NE061598_RS10395c (1800663-1801447)	x	++	NE061598_RS09345 (1799757-1800566)	MFS transporter
NE061598_RS09485	hypothetical protein	x	+	NE061598_RS10415c	+	x	NE061598_RS09470c	formate dehydrogenase

(1832656-1833129)				(1831782-1832566)			(1831242-1831661)	
NE061598_RS09560c (1845613-1846871)	MFS transporter	+	+	NE061598_RS10420c (1844680-1845464)	x	+	NE061598_RS09545 (1843687-1844582)	succinate semialdehyde dehydrogenase (NADP(+))
NE061598_RS09675 (1868129-1868682)	nicotinamide ribonucleoside (NR) uptake permease (PnuC) family protein	x	+	NE061598_RS10430 (1868706-1869490)	x	++	NE061598_RS09690c (1869729-1870478)	class I SAM-dependent methyltransferase

Table 6: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* A.II WY96 during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the WY96 chromosome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	Product		5'	Locus Tag (bp)	3'		Locus Tag (bp)	Product
FTW_RS00025c (4012-4565)	nicotinamide ribonucleoside uptake permease (PnuC)	x	+	FTW_RS00020c (3108-3988)	+	+	FTW_RS00015c (2692-3096)	hypothetical protein
FTW_RS00135 (25829-27087)	MFS transporter	+	++	FTW_RS09905 (27236-28116)	x	x	FTW_RS00150c (28117-28995)	MFS transporter
FTW_RS00270 (61081-61797)	formate dehydrogenase	++	+	FTW_RS09910 (61855-62735)	+	+	FTW_0058 (62739-63179)	formate dehydrogenase
FTW_RS09925 (92387-92885)	hypothetical protein	++	++	FTW_RS09935 (92947-93827)	x	+	FTW_RS00405 (93830-94339)	restriction endonuclease
FTW_RS00890 (187736-188479)	IS5/IS1182 family transposase	x	++	FTW_RS09950c (186754-187634)	+	+	FTW_RS00875 (186231-186776)	hypothetical protein
FTW_RS01060 (229118-230650)	glycerol-3-phosphate dehydrogenase/oxidase	++	++	FTW_RS09965c (227922-228802)	++	+++	FTW_0219 (226417-227925)	glycerol kinase
FTW_RS01275c (269914-270377)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	+	+	FTW_RS09980 (270407-271287)	+	+++	FTW_RS01290c (271407-273041)	molecular chaperone GroEL
FTW_RS01325 (276475-277125)	type I restriction endonuclease subunit R	x	++	FTW_RS0990c (275503-276383)	+	++	FTW_RS01310 (274692-275456)	peroxide stress protein YaaA

FTW_RS01455c (300838-301329)	N-acetyltransferase	x	++	FTW_RS10005c (299828-300708)	+	++	FTW_RS10000 (298664-299515)	hypothetical protein
FTW_RS01475c (305001-305231)	hypothetical protein	+	+	FTW_RS10010c (305309-306189)	x	+	FTW_RS01495 (306180-306419)	hypothetical protein
FTW_RS01565 (322449-323372)	hypothetical protein	++	+	FTW_RS10015c (321578-322458)	+	+	FTW_RS01550 (318743-321577)	hypothetical protein
FTW_RS01640 (332733-334247)	fumarate hydratase	+++	+	FTW_RS10020 (334732-335612)	+	+	FTW_RS01655c (335617-335982	IS1595 family transposase
FTW_RS01865c (374562-375113)	cyclopropane-fatty-acyl-phospholipid synthase	+	++	FTW_RS10025c (373658-374538)	+	+	FTW_RS01850c (373000-373686)	class I SAM-dependent methyltransferase
FTW_RS02080 (424583-425584)	dTDP-glucose 4,6-dehydratase	++	++	FTW_RS10030 (425713-426593)	+	+	FTW_RS02095c (426658-427173)	N-acetyltransferase
FTW_RS02220 (449354-449653)	transcriptional regulator	++	+	FTW_RS10035c (448316-449196)	x	++	FTW_RS02205c (446657-448063)	mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
FTW_RS02460 (495139-495408)	hypothetical protein	+	+	FTW_RS10055c (494268-495148)	+	++	FTW_RS02445 (491882-493918)	DNA ligase (NAD(+)) LigA
FTW_RS02580c (516560-516817)	hypothetical protein	x	+	FTW_RS10070 (516949-517829)	+	x	FTW_RS02595c (517830-518708)	MFS transporter
FTW_RS02825 (569522-570235)	pirin family protein	++	+	FTW_RS10075 (570259-571139)	x	+	FTW_RS02840c (571319-571588)	hypothetical protein
FTW_RS03060 (609243-609947)	DedA family protein	+	++	FTW_RS10085 (610239-611119)	+	++	FTW_RS03080 (611748-611936)	hypothetical protein
FTW_RS03350 (655628-657712)	glycine--tRNA ligase subunit beta	++	+	FTW_RS10105 (657852-658732)	+	+++	FTW_RS03365c (658928-659923)	hypothetical protein

FTW_RS03390 (661802-662366)	flavodoxin	x	+	FTW_RS10110 (660831-661711)	+	+	FTW_RS03370 (660111-660698)	IS5/IS1182 family transposase
FTW_RS03545c (694833-696029)	tRNA (Ile)-lysidine synthetase	++	+	FTW_RS10115c (693824-694704)	+	++	FTW_RS03530 (692706-693846)	MFS transporter
FTW_RS03645 (714061-714453)	hypothetical protein	++	++	FTW_RS10125 (714515-715395)	+	+++	FTW_RS03660 (715832-716440)	hypothetical protein
FTW_RS04040 (788856-790268)	dihydrolipoamide dehyrogenase	+++	+	FTW_RS10145 (790622-791502)	x	x	FTW_RS10150c (791541-791894)	hypothetical protein
FTW_RS10175 (935793-936047)	hypothetical protein	x	+	FTW_RS10170c (934922-935802)	+	++	FTW_RS04840 (933889-934944)	histidine-type phosphatase
FTW_RS05080c (977380-978522)	alpha-hydroxy-acid oxidizing protein	x	++	FTW_RS10195 (978617-979497)	+	++	FTW_RS05100c (979900-981090)	tyrosine--tRNA ligase
FTW_RS10225 (1115625-1115906)	hypothetical protein	+	++	FTW_RS10220c (1114754-1115634)	+++	+++	FTW_RS10215 (1114651-1114776)	hypothetical protein
FTW_RS06150c (1190603-1192051)	DUF3573 domain-containing protein	+	++	FTW_RS10240c (1189588-1190468)	+	+	FTW_RS06135c (1187663-1189643)	peptidase M13
FTW_RS06650c (1282205-1282579)	hypothetical protein	x	+	FTW_RS10250c (1281172-1282052)	+	++	FTW_RS06635c (1280397-1281044)	DNA mismatch repair protein MutT
FTW_RS06695c (1292622-1292993)	hypothetical protein	x	++	FTW_RS10260 (1293228-1294108)	+	x	FTW_RS06710c (1294422-1294631)	hypothetical protein
FTW_RS07025 (1353708-1354451)	IS5/IS1182 family transposase	x	+	FTW_RS10280 (1354599-1355479)	+	++	FTW_RS07040 (1355765-1356886)	polyamine ABC transporter ATP-binding protein
FTW_RS07075c (1361729-1362103)	hypothetical protein	x	+	FTW_RS10285c (1360787-1361667)	+	+++	FTW_RS07060 (1359596-1360591)	hypothetical protein

FTW_RS07150c (1376104-1378020)	ATP-dependent metallopeptidase	+++	+	FTW_RS10295c (1375002-1375882)	+	+++	FTW_RS07135c (1374207-1374929)	purine-nucleoside phosphorylase
FTW_RS07340 (1421446-1422234)	oxidoreductase	++	+	FTW_RS10305 (1422525-1423405)	x	+++	FTW_RS07355 (1424174-1426399)	catalase-peroxidase
FTW_RS07510 (1456385-1457209)	formamidopyrimidine-DNA glycosylase	+	+	FTW_RS10310 (1457280-1458160)	++	+++	FTW_RS07525c (1458168-1458716)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
FTW_RS07890c (1538167-1538358)	rRNA methyltransferase	x	+	FTW_RS10320 (1538349-1539229)	++	++	FTW_RS07905c (1539207-1539863)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
FTW_RS08200c (1608146-1608889)	IS5/IS1182 family transposase	+	++	FTW_RS10325 (1608991-1609871)	++	++	FTW_RS08215c (1609899-1610348)	nuclear transport factor 2 family protein
FTW_RS08280c (1623099-1624157)	hypothetical protein	x	+	FTW_RS10330 (1624428-1625308)	x	+	FTW_RS08300c (1625441-1626028)	IS5/IS1182 family transposase
FTW_RS08320 (1628455-1629327)	hypothetical protein	+	+	FTW_RS10335c (1627407-1628287)	+	+++	FTW_RS08305 (1626216-1627211)	hypothetical protein
FTW_RS08340c (1632442-1633524)	aminoacetone oxidase family FAD-binding enzyme	+	+	FTW_RS10345 (1633654-1634534)	+	+++	FTW_RS08355c (1634730-1635725)	hypothetical protein
FTW_RS08375 (1637675-1639561)	molecular chaperone HtpG	++	+	FTW_RS10350c (1636633-1637513)	x	+	FTW_RS08360 (1635913-1636500)	IS5/IS1182 family transposase
FTW_RS08390c (1640825-1641262)	50S ribosomal protein L17	+++	+	FTW_RS10355c (1639791-1640671)	+	++	FTW_RS08375 (1637675-1639561)	molecular chaperone HtpG
FTW_RS08805c (1704599-1705558)	hypothetical protein	++	+	FTW_RS10365c (1703229-1704109)	+	x	FTW_RS08780 (1702132-1703251)	MFS transporter
FTW_RS08965c (1730762-1730944)	hypothetical protein	+++	+	FTW_RS10370c (1729820-1730700)	+	+	FTW_RS08950c (1728786-1729817)	MFS transporter

FTW_RS09060 (1746529-1746723)	hypothetical protein	++	++	FTW_RS09065 (1746747-1747315)	+	x	FTW_RS09075 (1747389-1747760)	hypothetical protein
FTW_RS09110 (1757246-1759399)	primosomal protein N'	+	+	FTW_RS10385c (1756236-1757116)	++	+++	FTW_RS09095 (1754862-1756235)	2-isopropylmalate synthase
FTW_RS09335c (1805413-1805964)	hypothetical protein	x	+	FTW_RS10400c (1804509-1805389)	x	+	FTW_RS09320c (1803172-1804085)	peptidase
FTW_RS10410 (1827641-1827813)	endoglucanase	x	+	FTW_RS10405c (1826681-1827561)	+	+	FTW_RS09445 (1826089-1826703)	sodium:proton antiporter
FTW_RS09680c (1871123-1871458)	hypothetical protein	++	++	FTW_RS10425c (1870219-1871099)	x	+	FTW_RS09665 (1869322-1870217)	succinate-semialdehyde dehydrogenase (NADP(+))

Table 7: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* A.II W4114 during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the W4114 chromosome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
FT4114_RS00030 (4012-4565)	nicotinamide ribonucleoside uptake permease family protein	+	++	FT4114_RS09890c (3108-3988)	+	+	FT4114_RS00015c (2692-3096)	hypothetical protein
FT4114_RS00140 (25829-27086)	MFS transporter	+	+	FT4114_RS09900 (27235-28115)	+	X	FT4114_RS00155c (28116-28994)	MFS transporter
FT4114_RS00275 (61192-61908)	formate dehydrogenase	++	++	FT4114_RS09905 (61966-62846)	X	+	FT4114_RS00290 (62850-63290)	formate dehydrogenase
FT4114_RS00410 (93941-94450)	hypothetical protein	+	+	FT4114_RS09930 (93058-93938)	X	+	FT4114_RS00410 (93941-94450)	restriction endonuclease
FT4114_RS00895 (187847-188590)	IS5/IS1182 family transposase	X	++	FT4114_RS09945c (186865-187745)	+	++	FT4114_RS00880 (186342-186887)	hypothetical protein
FT4114_RS09965 (228903-229001)	rRNA methyltransferase	X	++	FT4114_RS09960c (228032-228912)	+	+++	FT4114_RS01055 (226528-228035)	glycerol kinase
FT4114_RS01285c (270023-270486)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	+	++	FT4114_RS09975 (270516-271396)	+	+++	FT4114_RS01300c (271516-273150)	molecular chaperone GroEL
FT4114_RS01335 (276584-277234)	type I restriction endonuclease subunit R	X	+	FT4114_RS09985c (275612-276492)	+	++	FT4114_RS01320 (274801-275565)	peroxide stress protein YaaA

FT4114_RS01465c (300947-301438)	N-acetyltransferase	X	X	FT4114_RS10000 (299937-300817)	+	+	FT4114_RS09995 (298773-299624)	hypothetical protein
FT4114_RS01505 (306289-306528)	hypothetical protein	++	+	FT4114_RS10005c (305418-306298)	X	X	FT4114_RS01485c (305110-305340)	hypothetical protein
FT4114_RS01575 (322558-323481)	hypothetical protein	++	++	FT4114_RS10010c (321687-322567)	+	+	FT4114_RS01565 (318852-321686)	hypothetical protein
FT4114_RS01650 (332842-334356)	fumarate hydratase	+++	++	FT4114_RS10015 (334839-335662)	+	+	FT4114_RS01665c (335667-336032)	IS1595 family transposase
FT4114_RS01875c (374612-375163)	cyclopropane-fatty-acyl-phospholipid synthase	X	++	FT4114_RS10020c (373708-374588)	X	X	FT4114_RS01860c (373050-373736)	class I SAM-dependent methyltransferase
FT4114_RS02090 (424625-425671)	dTDP-glucose 4, 6-dehydratase	++	++	FT4114_RS10025 (425800-426740)	+	+	FT4114_RS02105c (426805-427320)	N-acetyltransferase
FT4114_RS02230 (449501-449800)	transcriptional regulator	++	++	FT4114_RS10030c (448463-449343)	+	++	FT4114_RS02215c (446804-448210)	mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
FT4114_RS10050c (495508-495777)	hypothetical protein	X	++	FT4114_RS10050c (494637-495517)	+	X	FT4114_RS02460 (494434-494637)	hypothetical protein
FT4114_RS02595c (516929-517186)	hypothetical protein	X	++	FT4114_RS10065 (517334-518214)	+	X	FT4114_RS02610c (518215-519093)	MFS transporter
FT4114_RS02840 (569908-570621)	pirin family protein	++	++	FT4114_RS10070 (570645-571525)	+	+	FT4114_RS02855c (571705-571974)	hypothetical protein
FT4114_RS10075 (610422-610601)	peptidase M13	+++	+	FT4114_RS10080 (610625-611505)	+	+	FT4114_RS03090 (611477-611749)	phosphatase PAP2 family protein
FT4114_RS03365 (656014-658098)	glycerine--tRNA ligase subunit beta	++	+	FT4114_RS10100 (658238-659118)	X	+++	FT4114_RS03380c (659314-660309)	hypothetical protein

FT4114_RS03415c (665454-666197)	IS5/IS1182 family transposase	X	+	FT4114_RS10105c (664426-665306)	+	+++	FT4114_RS03400c (663019-664140)	polyamine ABC transporter ATP-binding protein
FT4114_RS03745 (726911-727282)	hypothetical protein	X	++	FT4114_RS10125c (725796-726676)	+	X	FT4114_RS03730 (725273-725482)	hypothetical protein
FT4114_RS03790 (737333-737707)	hypothetical protein	X	++	FT4114_RS10135 (737860-738740)	+	++	FT4114_RS03805 (738868-739515)	DNA mismatch repair protein MutT
FT4114_RS04290 (827861-829309)	DUF3573 domain-containing protein	+	++	FT4114_RS10145 (829444-830323)	+	++	FT4114_RS04305 (830929-831537)	hypothetical protein
FT4114_RS10160c (905350-905634)	IS1595 family transposase	+	++	FT4114_RS10165 (905684-906564)	X	X	FT4114_RS10170c (906603-906956)	hypothetical protein
FT4114_RS10195 (1050809-1051063)	hypothetical protein	+	+	FT4114_RS10190c (1049938-1050818)	+	++	FT4114_RS05485 (1048905-1049960)	histidine-type phosphatase
FT4114_RS05725c (1092434-1093570)	alpha-hydroxy-acid oxidizing protein	+	+	FT4114_RS10215 (1093665-1094485)	X	X	FT4114_RS05740c (1094773-1094952)	hypothetical protein
FT4114_RS10245 (1230675-1230956)	hypothetical protein	+	++	FT4114_RS10240c (12229804-1230684)	++	++	FT4114_RS10235 (1229701-1229826)	hypothetical protein
FT4114_RS06795c (1305652-1306848)	tRNA (Ile)-lysidine synthetase	++	++	FT4114_RS10260c (1304642-1305523)	+	+	FT4114_RS06780c (1302713-1304697)	peptidase M13
FT4114_RS06895 (1324880-1325272)	hypothetical protein	++	++	FT4114_RS10270 (1325334-1326214)	+	++	FT4114_RS06910c (1326192-1327332)	MFS transporter
FT4114_RS07050c (1357672-1358236)	flavodoxin	+	++	FT4114_RS10275 (1358327-1359206)	+	X	FT4114_RS07070c (1359341-1359997)	IS5/IS1182 family transposase
FT4114_RS07090c (1362289-1362663)	hypothetical protein	X	+	FT4114_RS10280c (1361364-1362227)	+	+++	FT4114_RS07080 (1360154-1361149)	hypothetical protein

FT4114_RS07165c (1376664-1378580)	ATP-dependent metallopeptidase HflB	+++	++	FT4114_RS10290c (1375562-1376442)	+	+++	FT4114_RS07150c (1374767-1375489)	purine-nucleoside phosphorylase
FT4114_RS10295c (1422962-1423095)	50S rRNA methyltransferase	++	+	FT4114_RS10300 (1423086-1423966)	+	+++	FT4114_RS07370 (1424735-1426960)	catalase-peroxidase
FT4114_RS07525 (1456934-1457758)	formamidopyrimidine-DNA glycosylase	+	+	FT4114_RS10305 (1457829-1458709)	+	+++	FT4114_RS07540c (1458717-1459265)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
FT4114_RS07910c (1538716-1538907)	rRNA methyltransferase	+	+	FT4114_RS10315 (1538898-1539778)	+	++	FT4114_RS07925c (1539756-1540412)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
FT4114_RS08220c (1608694-1609437)	IS5/IS1182 family transposase	X	++	FT4114_RS10320 (1609539-1610419)	++	++	FT4114_RS08235c (1610447-1610896)	nuclear transport factor 2 family protein
FT4114_RS08305c (1623647-1624705)	hypothetical protein	X	++	FT4114_RS10325 (1624976-1625856)	X	X	FT4114_RS08325c (1625989-1626642)	IS5/IS1182 family transposase
FT4114_RS08345 (1629037-1629909)	hypothetical protein	++	++	FT4114_RS10330c (1628106-1628869)	X	+++	FT4114_RS08330 (1626799-1627794)	hypothetical protein
FT4114_RS10335c (1634108-1634245)	hypothetical protein	+	++	FT4114_RS10340 (1634236-1635116)	+	+++	FT4114_RS08380c (1635312-1636307)	hypothetical protein
FT4114_RS08400 (1638292-1640178)	molecular chaperone HtpG	+++	+	FT4114_RS10345c (1637250-1638130)	+	X	FT4114_RS08385 (1636464-1637117)	IS5/IS1182 family transposase
FT4114_RS10355 (1641279-1641392)	rRNA methyltransferase	++	++	FT4114_RS10350c (1640408-1641288)	+	+++	FT4114_RS08400 (1638292-1640178)	molecular chaperone HtpG
FT4114_RS08825c (1704964-1705335)	hypothetical protein	+	++	FT4114_RS10360c (1703941-1704821)	+	+	FT4114_RS08805 (1702844-1703963)	MFS transporter
FT4114_RS10370 (1731356-1731430)	50S rRNA methyltransferase	+++	++	FT4114_RS10365c (1730485-1731365)	++	++	FT4114_RS08970c (1729451-1730482)	MFS transporter

FT4114_RS09085 (1747194-1747388)	hypothetical protein	++	++	FT4114_RS09090 (1747412-1747881)	+	X	FT4114_RS09100 (1748054-1748425)	hypothetical protein
FT4114_RS10385 (1757772-1757927)	hypothetical protein	+	++	FT4114_RS10380c (1756901-1757781)	++	+++	FT4114_RS09120 (1755527-1756900)	2-isopropylmalate synthase
FT4114_RS09360c (1806138-1806689)	hypothetical protein	X	+	FT4114_RS10395c (1805234-1806114)	+	++	FT4114_RS10390c (1804880-1805151)	hypothetical protein
FT4114_RS10405 (1828366-1828538)	endoglucanase	X	++	FT4114_RS10400c (1827406-1828286)	+	+	FT4114_RS09470 (1826814-1827428)	sodium:proton antiporter
FT4114_RS09705c (1871851-1872186)	hypothetical protein	++	+	FT4114_RS10420c (1870947-1871827)	X	++	FT4114_RS09690 (1870050-1870945)	succinate-semialdehyde (NADP (+))

Table 8: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* attenuated type B LVS during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the LVS genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
FTL_RS00030 (4207-4944)	NA+/H+ antiporter NHAP, fragment	++	+	FT_RS10080c (3336-4216)	+	+++	FTL_RS00015c (2921-3324)	symporter
FTL_RS00100c (17138-17472)	sugar transferase	+++	+	FTL_RS00095c (16234-17114)	+	++	FTL_RS00090 (15337-16232)	sugar transferase
FTL_RS00115c (20772-21387)	succinate-semialdehyde dehydrogenase	++	+	FTL_RS10085 (21378-22258)	+	+++	FTL_RS10095c (22726-23146)	endoglucanase
FTL_RS00285c (55312-55863)	hypothetical protein	++	+	FTL_RS10105c (54408-55288)	+	++	FTL_RS10100c (52734-54148)	beta-fructofuranosidase
FTL_RS00410c (77813-78498)	MFS transporter	++	+	FTL_RS10115c (76775-77604)	+	++	FTL_RS00395c (75931-76704)	IS5/IS1182 family transposase
FTL_RS00520 (99682-100287)	nicotinamide ribonucleoside (NR) uptake permease (PnuC) family protein	++	+	FTL_RS10125 (100311-101191)	+	x	FTL_RS00535c (101192-102064)	MFS transporter
FTL_RS00655 (129871-130587)	formate dehydrogenase	+++	+	FTL_RS10130 (130645-131525)	+	+++	FTL_RS00670c (131814-132894)	2-isopropylmalate synthase
FTL_RS00705c (141465-141835)	hypothetical protein	+++	+	FTL_RS10140c (140541-141322)	+	+	FTL_RS00690c (139997-140368)	hypothetical protein

FTL_RS00800 (158692-159012)	hypothetical protein	++	+	FTL_RS10145 (159036-159916)	+	+++	FTL_RS00815 (159888-160157)	hypothetical protein
FTL_RS10155 (168293-168367)	50S rRNA methyltransferase	x	+	FTL_RS10150c (167422-168302)	+	++	FTL_RS00865c (167216-167356)	hypothetical protein
FTL_RS00975 (184132-184326)	hypothetical protein	+	+	FTL_RS10160 (184350-185230)	+	++	FTL_RS00990c (185208-186326)	MFS transporter
FTL_RS01380 (247317-247754)	50s ribosomal protein L17	+++	+	FTL_RS10165 (247908-248734)	+	+	FTL_RS01390c (248743-249316)	IS5/IS1182 family transposase
FTL_RS01410 (251803-253689)	molecular chaperone HtpG	+++	+	FTL_RS10170c (250761-251641)	+	x	FTL_RS01395 (249570-250565)	hypothetical protein
FTL_RS01425c (254991-256340)	NADP-specific glutamate dehydrogenase	+++	+	FTL_RS10175c (253919-254799)	+	+++	FTL_RS01410 (251803-253689)	molecular chaperone HtpG
FTL_RS10185c (301298-301657)	hypothetical protein	+++	+	FTL_RS10190 (301648-302528)	+	++	FTL_RS01665 (302656-303303)	DNA mismatch repair protein MutT
FTL_RS01770 (320415-320867)	GCN5-related N-acetyltransferase	++	+	FTL_RS10200 (321013-321893)	+	+++	FTL_RS01785 (321853-322074)	TIGR03643 family protein
FTL_RS01855c (332031-332390)	DUF393cdomain-containing protein	+	+	FTL_RS10205c (331033-331913)	+	+	FTL_RS01840c (330636-331061)	DUF423 domain-containing protein
FTL_RS02165c (385950-386186)	hypothetical protein	+	+	FTL_RS10215c (385054-385880)	+	++	FTL_RS02155 (384311-385045)	IS5/IS1182 family transposase
FTL_RS02265 (406452-408107)	DUF3573 domain-containing protein	+++	+	FTL_RS10220 (408242-409071)	++	++	FTL_RS10225 (409142-409876)	IS5/IS1182 family transposase
FTL_RS10245 (489471-489620)	hypothetical protein	x	+	FTL_RS10240c (488545-489425)	+	x	FTL_RS02630 (487684-488544)	agmatine deiminase family protein

FTL_RS02680 (495122-495802)	hypothetical protein	x	+	FTL_RS10250c (493992-494872)	+	+++	FTL_RS02665 (493358-494014)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
FTL_RS02755 (505163-506677)	fumerate hydratase	+++	++	FTL_RS02760 (507162-507589)	++	x	FTL_RS02770 (507885-508169)	nuclease
FTL_RS03160 (594602-595603)	dTDP-D-glucose 4,6-dehydratase	+	+	FTL_RS10265 (595748-596628)	+	+++	FTL_RS03175 (596881-598287)	Mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase
FTL_RS03305 (620245-620544)	transcriptional regulator	+++	+	FTL_RS10270c (619191-620071)	+	++	FTL_RS03290 (618612-619126)	GNAT family N-acetyltransferase
FTL_RS10290 (666111-666200)	rRNA methyltransferase	x	+	FTL_RS10285c (665240-666120)	+	++	FTL_RS03535 (664885-665262)	hypothetical protein
FTL_RS10300 (670320-670421)	rRNA methyltransferase	++	+	FTL_RS10295c (669449-670329)	+	+++	FTL_RS03560c (668045-669166)	polyamine ABC transporter ATP-binding protein
FTL_RS03825 (722099-722491)	hypothetical protein	+++	+	FTL_RS10310 (722553-723433)	+	+++	FTL_RS03840 (723390-725036)	peptidase M13
FTL_RS10330 (769365-769442)	50S rRNA methyltransferase	x	+	FTL_RS10325c (768494-769374)	+	++	FTL_RS04045 (767192-768486)	DNA methyltransferase
FTL_RS04960 (938239-939381)	alpha-hydroxy-acid oxidizing protein	++	+	FTL_RS10380c (937264-938144)	+	++	FTL_RS04945 (936857-937036)	hypothetical protein
FTL_RS05065c (960473-960766)	tRNA threonylcarbamoyladenosine dehydratase	+++	+	FTL_RS10395 (960820-961700)	+	++	FTL_RS05080c (961721-961981)	heme biosynthesis protein HemY
FTL_RS05195c (978414-978668)	hypothetical protein	x	+	FTL_RS10400 (978659-979539)	+	+++	FTL_RS05210c (979517-980572)	histidine-type phosphatase
FTL_RS05750 (1073135-1073404)	hypothetical protein	x	+	FTL_RS10415c (1072264-1073144)	+	x	FTL_RS05735 (1071741-1071950)	hypothetical protein

FTL_RS05860c (1094545-1094802)	hypothetical protein	x	+	FTL_RS10425 (1094950-1095830)	+	+++	FTL_RS05875c (1095831-1096703)	MFS transporter
FTL_RS10435c (1143205-1143363)	hypothetical protein	+++	+	FTL_RS10440 (1143354-1144234)	+	++	FTL_RS06125 (1144261-1145118)	EamA/RhaT family transporter
FTL_RS06350c (1188996-1189184)	hypothetical protein	++	+	FTL_RS10450c (1188050-1188930)	+	+++	FTL_RS06335 (1186750-1187862)	3-deoxy-7-phosphoheptulonate synthase
FTL_RS10465 (1231746-1231883)	hypothetical protein	+	+	FTL_RS10460c (1230875-1231755)	+	x	FTL_RS06580c (1230409-1230903)	hypothetical protein
FTL_RS06615c (1236059-1236931)	hypothetical protein	+++	+	FTL_RS10480 (1237099-1237979)	x	+++	FTL_RS06630c (1238007-1238456)	nuclear transport factor 2 family protein
FTL_RS06710c (1253064-1253498)	hypothetical protein	x	+	FTL_RS10490 (1253489-1254369)	+	x	FTL_RS06725c (1254565-1255560)	hypothetical protein
FTL_RS06875c (1283273-1285357)	glycine--tRNA ligase subunit beta	+++	+	FTL_RS10510c (1282297-1283123)	+	++	FTL_RS06865 (1281554-1282288)	IS5/IS1182 family transposase
FTL_RS10045 (1300144-1300608)	hypothetical protein	+	+	FTL_RS10515c (1300731-1301611)	+	++	FTL_RS07005c (1301987-1302802)	DUF2797 domain-containing protein
FTL_RS07185c (1331929-1332633)	DedA family protein	++	+	FTL_RS10525c (1330757-1331637)	+	+	FTL_RS07165c (1330513-1330785)	phosphatase PAP2 family protein
FTL_RS07430c (1373395-1373769)	hypothetical protein	x	+	FTL_RS10540c (1372491-1373317)	+	++	FTL_RS07420 (1371748-1372482)	IS5/IS1182 family transposase
FTL_RS10560 (1387539-1387610)	50S rRNA methyltransferase	x	+	FTL_RS10555c (1386668-1387548)	+	+++	FTL_RS07490c (1385873-1386595)	purine-nucleoside phosphorylase
FTL_RS10580 (1438197-1438346)	hypothetical protein	x	+	FTL_RS10575c (1437326-1438206)	+	+++	FTL_RS07705c (1434341-1436566)	catalase-peroxidase

FTL_RS07890 (1473082-1473906)	Formamidopyrimidine-DNA glycosylase	++	+	FTL_RS10585 (1473977-1474857)	+	+++	FTL_RS07905c (1474865-1475413)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
FTL_RS08215c (1539150-1540106)	glycosyltransferase	+++	+	FTL_RS10605c (1538111-1538991)	+	++	FTL_RS08200 (1536066-1537826)	Dolichyl-phosphate-mannose--protein mannosyltransferase
FTL_RS08285c (1556471-1556662)	rRNA methyltransferase	++	+	FTL_RS10610 (1556653-1557434)	+	x	FTL_RS10615 (1557497-1557694)	hypothetical protein
FTL_RS08315c (1561074-1562108)	hypothetical protein	x	+	FTL_RS10620c (1560132-1561012)	+	x	FTL_RS08300 (1557758-1560128)	hypothetical protein
FTL_RS08375c (1573127-1573366)	hypothetical protein	++	+	FTL_RS10625 (1573357-1574237)	+	+++	FTL_RS08390c (1574234-1575742)	glycerol kinase
FTL_RS10640c (1614741-1614812)	50S rRNA methyltransferase	x	+	FTL_RS10645 (1614803-1615683)	x	+	FTL_RS08565 (1615756-1615986)	hypothetical protein
FTL_RS08585 (1619653-1620144)	N-acetyltransferase	+	+	FTL_RS10650 (1620290-1621170)	+	++	FTL_RS08600c (1621483-1622334)	hypothetical protein
FTL_RS08715c (1643645-1644409)	type I restriction endonuclease subunit R	x	+	FTL_RS10655 (1644403-1645283)	+	+++	FTL_RS08730c (1645403-1647037)	molecular chaperone GroEL
FTL_RS08765 (1650405-1650866)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	++	+	FTL_RS10660c (1649495-1650375)	+	+++	FTL_RS08750 (1648680-1649444)	peroxide stress protein YaaA
FTL_RS10675c (1691889-1691987)	rRNA methyltransferase	++	+	FTL_RS10680 (1691978-1692858)	+	+++	FTL_RS09000c (1693508-1694548)	hypothetical protein
FTL_RS10715 (1749734-1749811)	50S rRNA methyltransferase	+++	+	FTL_RS10710c (1748863-1749743)	+	++	FTL_RS09280c (1747691-1748830)	ATP-binding protein
FTL_RS10755 (1819446-1819571)	rRNA methyltransferase	x	++	FTL_RS10750c (1818626-1819455)	+	++	FTL_RS10060c (1817821-1818555)	IS5/IS1182 family transposase

FTL_RS10760c (1822688-1822900)	hypothetical protein	+++	+	FTL_RS10765 (1822946-1823826)	+	x	FTL_RS09640c (1824022-1825017)	hypothetical protein
FTL_RS09805c (1858202-1858804)	IS5/IS1182 family transposase	++	+	FTL_RS10775c (1857257-1858137)	+	++	FTL_RS09790c (1856902-1857174)	hypothetical protein
FTL_RS10780c (1873217-1873372)	hypothetical protein	+	+	FTL_RS10785 (1873363-1874243)	+	+++	FTL_RS09890c (1874376-1875136)	SAM-dependent methyltransferase

Table 9: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* subsp. *novicida* U112 during mid exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to the IS*Ftu1* gene in the U112 genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
FTN_RS09355c (1721368-1721619)	IS1595 family transposase	X	++	FTN_RS09360 (1721669-1722536)	+	X	FTN_RS08280c (1722564-1723307)	IS5 family transposase

Table 10: Expression of the 5' and 3' IS*Ftu1* regions and the adjacent genes in *F. tularensis* A.I strain NE061598 treated with SNAP during early exponential phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the NE061598 genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
NE061598_RS09860 (3957-4178)	hypothetical protein	+	+	NE061598_RS00020c (3204-3988)	+	+	NE061598_RS00015c (2692-3096)	hypothetical protein
NE061598_RS00040 (8039-8374)	hypothetical protein	++	+	NE061598_RS00045 (8446-8915)	+	+	NE061598_RS00055 (9017-9526)	restriction endonuclease
NE061598_RS00535 (103197-103940)	IS5 family transposase IS*Ftu2*	x	+	NE061598_RS09885c (102311-103095)	x	+	NE061598_RS00520 (101692-102237)	hypothetical protein
NE061598_RS09905 (144438-144536)	rRNA methyltransferase	+	++	NE061598_RS09900c (143663-144447)	+	+++	NE061598_RS00695 (142062-143570)	glycerol kinase
NE061598_RS00925c (185593-186057)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	+	+	NE061598_RS09910 (186087-186871)	+	++	NE061598_RS00940 (187050-187322)	hypothetical protein
NE061598_RS09920c (234207-234362)	hypothetical protein	+	++	NE061598_RS09925 (234353-235137)	+	+	NE061598_RS01180 (235205-235474)	hypothetical protein
NE061598_RS01240c (243642-243836)	hypothetical protein	++	++	NE061598_RS09930c (242834-243618)	+	+	NE061598_RS01225c (241704-242735)	MFS transporter
NE061598_RS09935c (259599-259673)	50S rRNA methyltransferase	+++	++	NE061598_RS09940 (259664-260448)	+	x	NE061598_RS01350 (260618-260989)	hypothetical protein
NE061598_RS01385c	hypothetical protein	+	+	NE061598_RS09950c	+	+++	NE061598_RS01370	2-isopropylmalate synthase

(270488-270859)				(269561-270345)			(268091-269464)	
NE061598_RS01480 (287318-287638)	hypothetical protein	+++	++	NE061598_RS09960 (287662-288446)	+	+	NE061598_RS01500c (288520-289639)	MFS transporter
NE061598_RS01890 (350683-351120)	50S ribosomal protein L17	+++	++	NE061598_RS09965 (351274-352058)	+	+	NE061598_RS01905c (352132-352726)	IS5/IS1182 family transposase
NE061598_RS01930 (355425-356483)	hypothetical protein	x	+	NE061598_RS09970c (354274-355058)	+	+++	NE061598_RS01910 (352987-353982)	hypothetical protein
NE061598_RS02005 (370618-371490)	hypothetical protein	+	++	NE061598_RS09985c (369666-370450)	+	++	NE061598_RS01990 (369093-369542)	nuclear transport factor 2 family protein
NE061598_RS09990c (375691-375828)	hypothetical protein	+	+	NE061598_RS09995 (375819-376603)	+	++	NE061598_RS02040c (376927-378813)	molecular chaperone HtpG
NE061598_RS02040c (376927-378813)	molecular chaperone HtpG	++	++	NE061598_RS10000 (378975-379759)	+	+++	NE061598_RS02055c (380051-381046)	hypothetical protein
NE061598_RS02375c (454399-454674)	hypothetical protein	x	+	NE061598_RS10010c (453417-454201)	+	++	NE061598_RS02360 (452687-453343)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
NE061598_RS02750 (535404-536503)	alpha-hydroxy-acid oxidizing enzyme	+	+	NE061598_RS10025c (534525-535309)	+	+++	NE061598_RS02735 (533873-534421)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
NE061598_RS02995c (578188-578442)	hypothetical protein	x	++	NE061598_RS10045 (578433-579217)	+	+++	NE061598_RS03010 (579599-580720)	polyamine ABC transporter ATP-binding protein
NE061598_RS03700c (711649-712473)	formamidopyrimidine-DNA glycosylase	+	++	NE061598_RS10090c (710794-711578)	+	++	NE061598_RS03685 (709103-710293)	tyrosine--tRNA ligase
NE061598_RS10100 (745121-745270)	hypothetical protein	+	+	NE061598_RS10095c (744346-745130)	x	+++	NE061598_RS03850c (741257-743482)	catalase-peroxidase
NE061598_RS04100		++	+	NE061598_RS10110	+	+++	NE061598_RS04115	

(782835-784919)	glycine--tRNA ligase subunit beta			(785059-785843)			(786012-786734)	purine-nucleoside phosphorylase
NE061598_RS10120c (799182-799301)	rRNA methyltransferase	+	+	NE061598_RS10125 (799292-800076)	x	+	NE061598_RS10130 (800121-800363)	hypothetical protein
NE061598_RS10140 (839071-839250)	peptidase M13	+++	++	NE061598_RS10145 (839274-840058)	+	++	NE061598_RS04435 (840282-840929)	DNA mismatch repair protein MutT
NE061598_RS04930 (929421-930866)	DUF3573 domain-containing protein	+	+	NE061598_RS10150 (931001-931785)	+	x	NE061598_RS04945c (932093-932272)	hypothetical protein
NE061598_RS05735c (1074428-1074802)	hypothetical protein	x	+	NE061598_RS10170c (1073475-1073891)	x	++	NE061598_RS05720 (1072346-1072936)	histidine-type phosphatase
NE061598_RS05780c (1084894-1085265)	hypothetical protein	x	++	NE061598_RS10175 (1085499-1086283)	+	+	NE061598_RS05795c (1086693-1086902)	hypothetical protein
NE061598_RS10185 (1153765-1153890)	hypothetical protein	+++	++	NE061598_RS10190c (1153964-1154748)	+	+	NE061598_RS10195 (1154739-1155020)	hypothetical protein
NE061598_RS06510c (1230848-1231240)	hypothetical protein	++	++	NE061598_RS10210c (1230002-1230786)	+	++	NE061598_RS06495c (1227979-1229961)	peptidase M13
NE061598_RS06610 (1249232-1250428)	tRNA(Ile)-lysidine synthetase	++	+	NE061598_RS10220 (1250573-1251357)	+	+	NE061598_RS06625c (1251431-1252759)	MFS transporter
NE061598_RS10230c (1283681-1283875)	hypothetical protein	x	+	NE061598_RS10235 (1283869-1284653)	x	+++	NE061598_RS06790c (1284945-1285940)	hypothetical protein
NE061598_RS06810c (1287693-1288406)	pirin family protein	++	++	NE061598_RS10240c (1286869-1287653)	+	+	NE061598_RS06795 (1286201-1286795)	IS5/IS1182 family transposase
NE061598_RS07055 (1341129-1341386)	hypothetical protein	x	+	NE061598_RS10245c (1340197-1340981)	+	+	NE061598_RS07040 (1339291-1340100)	MFS transporter
NE061598_RS07165c	hypothetical protein	x	++	NE061598_RS10255	+	++	NE061598_RS10260c	hypothetical protein

(1362500-1362769)				(1362760-1363544)			(1363614-1363993)	
NE061598_RS07405c (1408369-1408668)	transcriptional regulator	+	+	NE061598_RS10275 (1408843-1409627)	x	+	NE061598_RS07420c (1409788-1410303)	N-acetyltransferase
NE061598_RS10295 (1465324-1465683)	hypothetical protein	++	+	NE061598_RS10290c (1464549-1465333)	x	x	NE061598_RS07700c (1464209-1464481)	phosphatase PAP2 family protein
NE061598_RS08060 (1542117-1542599)	cyclopropane-fatty-acyl-phospholipid synthase	+	+	NE061598_RS10310 (1542623-1543407)	x	+	NE061598_RS08075 (1543475-1544161)	class I SAM-dependent methyltransferase
NE061598_RS08285 (1582443-1582682)	hypothetical protein	+	+	NE061598_RS10315c (1581668-1582452)	+	+	NE061598_RS08265 (1581193-1581661)	IS1595 family transposase
NE061598_RS08360c (1599211-1600725)	fumarate hydratase	+++	++	NE061598_RS10320c (1597942-1598726)	+	+	NE061598_RS08345 (1595011-1597845)	hypothetical protein
NE061598_RS08465 (1615614-1616105)	GNAT family N-acetyltransferase	x	+	NE061598_RS10335 (1616251-1617035)	+	++	NE061598_RS08480c (1617444-1618295)	hypothetical protein
NE061598_RS08590c (1639747-1640511)	type I restriction endonuclease subunit R	+	++	NE061598_RS10345 (1640505-1641289)	+	+	08605NE061598_RS (1641645-1643069)	beta-fructofuranosidase
NE061598_RS10355c (1671285-1671456)	endoglucanase	x	+	NE061598_RS10360 (1671536-1672320)	+	+	NE061598_RS08740c (1672394-1673008)	sodium:proton antiporter
NE061598_RS08850 (1693185-1693709)	hypothetical protein	x	+	NE061598_RS10365 (1693733-1694517)	+	++	NE061598_RS08865c (1694660-1695454)	peroxide stress protein YaaA
NE061598_RS08900c (1699767-1700483)	formate dehydrogenase	++	+	NE061598_RS10375c (1698925-1699709)	x	+++	NE061598_RS08885 (1697075-1698709)	molecular chaperone GroEL
NE061598_RS10400 (1801438-1801557)	hypothetical protein	++	++	NE061598_RS10395c (1800663-1801447)	+	x	NE061598_RS09345 (1799757-1800566)	MFS transporter
NE061598_RS09485	hypothetical protein	x	+	NE061598_RS10415c	+	+	NE061598_RS09470c	formate dehydrogenase

(1832656-1833129)				(1831782-1832566)			(1831242-1831661)	
NE061598_RS09560c (1845613-1846871)	MFS transporter	+	++	NE061598_RS10420c (1844680-1845464)	+	+	NE061598_RS09545 (1843687-1844582)	succinate semialdehyde dehydrogenase (NADP(+))
NE061598_RS09675 (1868129-1868682)	nicotinamide ribonucleoside (NR) uptake permease (PnuC) family protein	x	+	NE061598_RS10430 (1868706-1869490)	+	++	NE061598_RS09690c (1869729-1870478)	class I SAM-dependent methyltransferase

Table 11: Expression of the 5' and 3' IS*Ftu1* regions and adjacent genes in *F. tularensis* A.I strain NE061598 treated with spermine during early exponential growth phase. Locus tag, base pair location, and proposed function of each gene adjacent to each IS*Ftu1* gene in the NE061598 genome are listed. A locus tag with "c" at the end indicates the complement sequence. The RNA expression levels for IS*Ftu1* and the adjacent genes are shown in the columns next to the respective locus tag. The notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

5' Adjacent Gene			IS*Ftu1*			3' Adjacent Gene		
Locus Tag (bp)	**Product**		**5'**	**Locus Tag (bp)**	**3'**		**Locus Tag (bp)**	**Product**
NE061598_RS09860 (3957-4178)	hypothetical protein	+	+	NE061598_RS00020c (3204-3988)	+	++	NE061598_RS00015c (2692-3096)	hypothetical protein
NE061598_RS00040 (8039-8374)	hypothetical protein	++	+	NE061598_RS00045 (8446-8915)	+	++	NE061598_RS00055 (9017-9526)	restriction endonuclease
NE061598_RS00535 (103197-103940)	IS5 family transposase IS*Ftu2*	x	+	NE061598_RS09885c (102311-103095)	x	+	NE061598_RS00520 (101692-102237)	hypothetical protein
NE061598_RS09905 (144438-144536)	rRNA methyltransferase	x	++	NE061598_RS09900c (143663-144447)	x	+++	NE061598_RS00695 (142062-143570)	glycerol kinase
NE061598_RS00925c (185593-186057)	23S rRNA (pseudouridine (1915)-N(3))-methyltransferase RlmH	+	+	NE061598_RS09910 (186087-186871)	+	++	NE061598_RS00940 (187050-187322)	hypothetical protein
NE061598_RS09920c (234207-234362)	hypothetical protein	+	+	NE061598_RS09925 (234353-235137)	+	+	NE061598_RS01180 (235205-235474)	hypothetical protein
NE061598_RS01240c (243642-243836)	hypothetical protein	++	++	NE061598_RS09930c (242834-243618)	+	+	NE061598_RS01225c (241704-242735)	MFS transporter
NE061598_RS09935c (259599-259673)	50S rRNA methyltransferase	+++	+++	NE061598_RS09940 (259664-260448)	+	x	NE061598_RS01350 (260618-260989)	hypothetical protein
NE061598_RS01385c	hypothetical protein	x	+	NE061598_RS09950c	+	+++	NE061598_RS01370	2-isopropylmalate synthase

(270488-270859)				(269561-270345)			(268091-269464)	
NE061598_RS01480 (287318-287638)	hypothetical protein	+++	++	NE061598_RS09960 (287662-288446)	x	++	NE061598_RS01500c (288520-289639)	MFS transporter
NE061598_RS01890 (350683-351120)	50S ribosomal protein L17	+++	+	NE061598_RS09965 (351274-352058)	+	x	NE061598_RS01905c (352132-352726)	IS5/IS1182 family transposase
NE061598_RS01930 (355425-356483)	hypothetical protein	x	+	NE061598_RS09970c (354274-355058)	+	+++	NE061598_RS01910 (352987-353982)	hypothetical protein
NE061598_RS02005 (370618-371490)	hypothetical protein	x	++	NE061598_RS09985c (369666-370450)	+	++	NE061598_RS01990 (369093-369542)	nuclear transport factor 2 family protein
NE061598_RS09990c (375691-375828)	hypothetical protein	+	++	NE061598_RS09995 (375819-376603)	x	++	NE061598_RS02040c (376927-378813)	molecular chaperone HtpG
NE061598_RS02040c (376927-378813)	molecular chaperone HtpG	++	+	NE061598_RS10000 (378975-379759)	+	+++	NE061598_RS02055c (380051-381046)	hypothetical protein
NE061598_RS02375c (454399-454674)	hypothetical protein	++	++	NE061598_RS10010c (453417-454201)	+	++	NE061598_RS02360 (452687-453343)	23S rRNA (guanosine (2251)-2'-O)-methyltransferase RlmB
NE061598_RS02750 (535404-536503)	alpha-hydroxy-acid oxidizing enzyme	+	+	NE061598_RS10025c (534525-535309)	+	+++	NE061598_RS02735 (533873-534421)	pyridoxal 5'-phosphate synthase glutaminase subunit PdxT
NE061598_RS02995c (578188-578442)	hypothetical protein	x	+	NE061598_RS10045 (578433-579217)	x	+++	NE061598_RS03010 (579599-580720)	polyamine ABC transporter ATP-binding protein
NE061598_RS03700c (711649-712473)	formamidopyrimidine-DNA glycosylase	+	+	NE061598_RS10090c (710794-711578)	+	++	NE061598_RS03685 (709103-710293)	tyrosine--tRNA ligase
NE061598_RS10100 (745121-745270)	hypothetical protein	+	+	NE061598_RS10095c (744346-745130)	+	+++	NE061598_RS03850c (741257-743482)	catalase-peroxidase
NE061598_RS04100		++	+	NE061598_RS10110	x	+++	NE061598_RS04115	

(782835-784919)	glycine--tRNA ligase subunit beta			(785059-785843)			(786012-786734)	purine-nucleoside phosphorylase
NE061598_RS10120c (799182-799301)	rRNA methyltransferase	+	+	NE061598_RS10125 (799292-800076)	+	x	NE061598_RS10130 (800121-800363)	hypothetical protein
NE061598_RS10140 (839071-839250)	peptidase M13	+++	++	NE061598_RS10145 (839274-840058)	+	++	NE061598_RS04435 (840282-840929)	DNA mismatch repair protein MutT
NE061598_RS04930 (929421-930866)	DUF3573 domain-containing protein	++	+	NE061598_RS10150 (931001-931785)	+	x	NE061598_RS04945c (932093-932272)	hypothetical protein
NE061598_RS05735c (1074428-1074802)	hypothetical protein	x	+	NE061598_RS10170c (1073475-1073891)	x	+	NE061598_RS05720 (1072346-1072936)	histidine-type phosphatase
NE061598_RS05780c (1084894-1085265)	hypothetical protein	x	+	NE061598_RS10175 (1085499-1086283)	+	x	NE061598_RS05795c (1086693-1086902)	hypothetical protein
NE061598_RS10185 (1153765-1153890)	hypothetical protein	+++	+	NE061598_RS10190c (1153964-1154748)	x	x	NE061598_RS10195 (1154739-1155020)	hypothetical protein
NE061598_RS06510c (1230848-1231240)	hypothetical protein	++	+	NE061598_RS10210c (1230002-1230786)	+	++	NE061598_RS06495c (1227979-1229961)	peptidase M13
NE061598_RS06610 (1249232-1250428)	tRNA(Ile)-lysidine synthetase	++	+	NE061598_RS10220 (1250573-1251357)	+	++	NE061598_RS06625c (1251431-1252759)	MFS transporter
NE061598_RS10230c (1283681-1283875)	hypothetical protein	x	+	NE061598_RS10235 (1283869-1284653)	x	+++	NE061598_RS06790c (1284945-1285940)	hypothetical protein
NE061598_RS06810c (1287693-1288406)	pirin family protein	++	+	NE061598_RS10240c (1286869-1287653)	x	x	NE061598_RS06795 (1286201-1286795)	IS5/IS1182 family transposase
NE061598_RS07055 (1341129-1341386)	hypothetical protein	x	++	NE061598_RS10245c (1340197-1340981)	+	x	NE061598_RS07040 (1339291-1340100)	MFS transporter
NE061598_RS07165c	hypothetical protein	x	+	NE061598_RS10255	+	++	NE061598_RS10260c	hypothetical protein

(1362500-1362769)				(1362760-1363544)			(1363614-1363993)	
NE061598_RS07405c (1408369-1408668)	transcriptional regulator	x	+	NE061598_RS10275 (1408843-1409627)	+	+	NE061598_RS07420c (1409788-1410303)	N-acetyltransferase
NE061598_RS10295 (1465324-1465683)	hypothetical protein	+	++	NE061598_RS10290c (1464549-1465333)	+	+	NE061598_RS07700c (1464209-1464481)	phosphatase PAP2 family protein
NE061598_RS08060 (1542117-1542599)	cyclopropane-fatty-acyl-phospholipid synthase	x	+	NE061598_RS10310 (1542623-1543407)	x	+	NE061598_RS08075 (1543475-1544161)	class I SAM-dependent methyltransferase
NE061598_RS08285 (1582443-1582682)	hypothetical protein	+	+	NE061598_RS10315c (1581668-1582452)	+	+	NE061598_RS08265 (1581193-1581661)	IS1595 family transposase
NE061598_RS08360c (1599211-1600725)	fumarate hydratase	+++	+	NE061598_RS10320c (1597942-1598726)	+	+	NE061598_RS08345 (1595011-1597845)	hypothetical protein
NE061598_RS08465 (1615614-1616105)	GNAT family N-acetyltransferase	x	+	NE061598_RS10335 (1616251-1617035)	+	++	NE061598_RS08480c (1617444-1618295)	hypothetical protein
NE061598_RS08590c (1639747-1640511)	type I restriction endonuclease subunit R	+	++	NE061598_RS10345 (1640505-1641289)	+	+	08605NE061598_RS (1641645-1643069)	beta-fructofuranosidase
NE061598_RS10355c (1671285-1671456)	endoglucanase	+	+	NE061598_RS10360 (1671536-1672320)	+	+	NE061598_RS08740c (1672394-1673008)	sodium:proton antiporter
NE061598_RS08850 (1693185-1693709)	hypothetical protein	x	+	NE061598_RS10365 (1693733-1694517)	+	++	NE061598_RS08865c (1694660-1695454)	peroxide stress protein YaaA
NE061598_RS08900c (1699767-1700483)	formate dehydrogenase	++	+	NE061598_RS10375c (1698925-1699709)	x	+++	NE061598_RS08885 (1697075-1698709)	molecular chaperone GroEL
NE061598_RS10400 (1801438-1801557)	hypothetical protein	++	+	NE061598_RS10395c (1800663-1801447)	x	x	NE061598_RS09345 (1799757-1800566)	MFS transporter
NE061598_RS09485	hypothetical protein	x	++	NE061598_RS10415c	+	+	NE061598_RS09470c	formate dehydrogenase

(1832656-1833129)				(1831782-1832566)			(1831242-1831661)	
NE061598_RS09560c (1845613-1846871)	MFS transporter	+	++	NE061598_RS10420c (1844680-1845464)	+	++	NE061598_RS09545 (1843687-1844582)	succinate semialdehyde dehydrogenase (NADP(+))
NE061598_RS09675 (1868129-1868682)	nicotinamide ribonucleoside (NR) uptake permease (PnuC) family protein	x	+	NE061598_RS10430 (1868706-1869490)	+	++	NE061598_RS09690c (1869729-1870478)	class I SAM-dependent methyltransferase

Table 12: Summary of expression observed for the 5' and 3' IS*Ftu1* regions and adjacent genes in *F. tularensis* strains. The expression levels for the 5' and 3 regions of ISFtu1 and adjacent genes in F. tularensis Schu S4, NE061598 (without and with SNAP or spermine), WY96, W4114, and LVS were obtained from the results shown in Tables 4-11 and are summarized below. The overall number of expressed adjacent genes of regions of IS*Ftu1* is shown in the last column, and the notation "x" indicates no expression, "+" indicates low expression, "++" indicates moderate expression, while a "+++" indicates high expression.

***F. tularensis* Strain and Region**	**x**	**+**	**++**	**+++**	**Total Expressed**
Schu S4 5' Adjacent Gene	20	13	12	5	30
Schu S4 5' Region of IS*Ftu1*	2	36	10	2	48
Schu S4 3' Region of IS*Ftu1*	12	36	2	0	38
Schu S4 3' Adjacent Gene	7	18	15	10	43
NE061598 5' Adjacent Gene	28	10	7	2	19
NE061598 5' Region of IS*Ftu1*	2	44	1	0	45
NE061598 3' Region of IS*Ftu1*	32	15	0	0	15
NE061598 3' Adjacent Gene	18	10	10	9	29
WY96 5' Adjacent Gene	16	13	15	5	33
WY96 5' Region of IS*Ftu1*	0	32	17	0	49
WY96 3' Region of IS*Ftu1*	11	32	5	1	38
WY96 3' Adjacent Gene	6	18	13	12	43
W4114 5' Adjacent Gene	15	15	14	5	34
W4114 5' Region of IS*Ftu1*	1	16	32	0	48
W4114 3' Region of IS*Ftu1*	10	35	4	0	39
W4114 3' Adjacent Gene	12	12	13	12	37
LVS 5' Adjacent Gene	16	18	17	8	43
LVS 5' Region of IS*Ftu1*	0	57	2	0	59
LVS 3' Region of IS*Ftu1*	2	55	2	0	57
LVS 3' Adjacent Gene	10	5	22	22	49
NE061598+SNAP 5' Adjacent Gene	14	17	10	6	33
NE061598+SNAP 5' Region of IS*Ftu1*	0	27	20	0	47
NE061598+SNAP 3' Region of IS*Ftu1*	9	38	0	0	38
NE061598+SNAP 3' Adjacent Gene	4	21	12	10	43
NE061598+Spermine 5' Adjacent Gene	17	13	11	6	30
NE061598+Spermine 5' Region of IS*Ftu1*	0	34	12	1	47
NE061598+Spermine 3' Region of IS*Ftu1*	13	34	0	0	34
NE061598+Spermine 3' Adjacent Gene	9	12	16	10	38

Genes Co-transcribed with IS*Ftu1*

Analysis of RNA-Seq transcriptome data from *F. tularensis* A.I strain Schu S4 revealed numerous genes that appeared to be co-transcribed with 50 IS*Ftu1* ORFs (Table 4). Based on these initial analyses and the annotation for Schu S4, we focused on several moderately or highly transcribed genes that were adjacent to a moderately or highly expressed IS*Ftu1* gene. More specifically, IS*Ftu1* locus tags FTT_0245, FTT_0274, and FTT_0823 were selected for further assessment, and the genetic organization, as well as flanking genes are shown in Figure 11.

Located 360 bp upstream of the highly expressed IS*Ftu1* FTT_0247 is locus tag FTT_0245 (Figure 11A). This gene is predicted to encode a universal stress protein (Usp) in *F. tularensis* Schu S4. Usp is a small cytoplasmic bacterial protein whose expression is enhanced during exposure to stress agents and presumably enhances the rate of cell survival, especially during prolonged exposures. The 278 amino acid Usp in *F. tularensis* Schu S4 shared the highest similarity to Usp in *Methanosarcina maizae* (323 residues) with 31.85% amino acid identity. Between the *usp* gene (FTT_0245) and IS*Ftu1* FTT_0247 is hypothetical pseudogene (FTT_0246c), and downstream of this IS element gene is the locus tag FTT_0248, which is predicted to encode a hypothetical protein followed by the *feoB* gene (FTT_0249). The *feoB* gene encodes a ferrous iron transport protein that shares the highest amino acid identity to FeoB in *Piscirickettsia salmonis* and may supply iron to *F. tularensis* under anaerobic conditions, similar to the function of this integral membrane protein in *E. coli* (Kammler et al., 1993).

The locus tag FTT_t07 encodes a tRNA-Gln and is 439 bp upstream of expressed IS*Ftu1* FTT_0273 (Figure 11B). Also upstream and directly adjacent to this IS*Ftu1* is the highly expressed locus tag FTT_0272, which is predicted to encode a hypothetical protein.

FTT_0274 is located 114 bp downstream of the expressed IS*Ftu1* FTT_0273 and was predicted to encode a hypothetical membrane protein in earlier versions of the Schu S4 annotation by NCBI, but is absent in the newer rendition, and the next downstream locus tag (FTT_0275c) is now identified as a pseudogene.

Located 1038 bp upstream of expressed IS*Ftu1* FTT_0822 is rp1T, a 50S ribosomal protein L20, followed by FTT_0821, a hypothetical protein (Figure 11C). Locus tag FTT_0823 is located 224 bp downstream of the expressed IS*Ftu1* FTT_0822 and is predicted to encode a MutT protein. MutT is considered a mutator since this protein has been determined to increase the rate of AT:GC transversions in other bacterial species, and is a 7,8-dihydro-8-oxoguanine triphosphatase that converts 8-oxoguanine triphosphatase to the monophosphate (Akiyama et al., 1987; Tajiri et al., 1995). In E. coli, MutT degrades oxidatively damaged guanine (Maki and Sekiguchi, 1992).

To confirm transcriptional co-expression of the IS*Ftu1* ORF and the adjacent gene, RT-PCR was performed with DNase-treated RNA isolated from mid exponential growth phase *F. tularensis* A.I strains Schu S4 and NE061598. Following cDNA synthesis, genes upstream and downstream of IS*Ftu1* were amplified with the appropriate primer pair shown in Table 2. All of the RT-PCR reactions produced the expected amplicons for each region that was co-transcribed with an IS*Ftu1*, and the positive control using *F. tularensis* genomic DNA as the template produced the expected 144 bp amplicon (Figure 12).

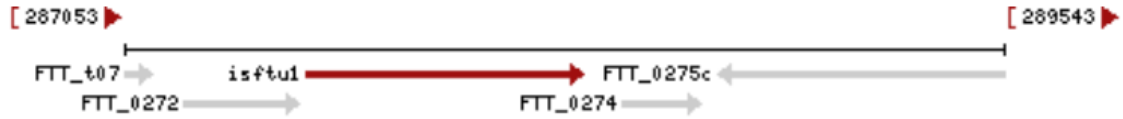
In more detail, for both *F. tularensis* Schu S4 and NE061568 the FTT_0245 specific forward primer and IS*Ftu1* specific reverse primer produced a 113 bp amplicon, and the IS*Ftu1* specific forward primer and FTT_0248 specific reverse primer produced a 126 bp amplicon for the first region of interest. The FTT_0272 specific forward primer and IS*Ftu1* specific reverse primer produced a 118 bp amplicon, and the IS*Ftu1* 5’ specific forward primer and FTT_0274 specific reverse primer produced a 106 bp amplicon for the

second region of interest. Finally, the FTT_0821 specific forward primer and IS*Ftu1* specific reverse primer produced a 111 bp amplicon, and the IS*Ftu1* specific forward primer and FTT_0823 specific reverse primer produced a 143 bp amplicon for the third region of interest.

Although the RT-PCR products produced in both Schu S4 and NE061598 for the 3' region of an IS*Ftu1* gene and the associated 5' region of the downstream genes identified as FTT_0248 and FTT_0274 were more abundant, unlike the similar abundance observed for the co-transcribed 5' and 3' adjacent genes of IS*Ftu1* FTT_0822, qRT-PCR is needed to confirm these findings. Together these results confirmed that either the 5' or 3' region of IS*Ftu1* was co-transcribed with an adjacent gene.

A

B



C

Figure 11: Genetic organization highly expressed IS*Ftu1* genes in *F. tularensis* A.I Schu S4. Directionality of highly expressed **(A)** IS*Ftu1* FTT_0247 **(B)** IS*Ftu1* FTT_0273, and **(C)** IS*Ftu1* FTT_0822 genes (red arrows), along with surrounding genes in *F. tularensis* A.I Schu S4 are shown. The beginning and ending nucleotide position of the loci is denoted at the top left and top right for the associated region in the chromosome, respectively. Images were obtained from NCBI at URL http://www.ncbi.nlm.nih.gov/gene/. Of note, FTT_0274 is no longer identified in the newest NCBI annotation of the Schu S4 genome, and the *usp*, *feoB*, rplT, and *arsC1* genes shown above are also identified as locus tags FTT_0245, FTT_0249, FTT_0820, and FTT_0824, respectively.

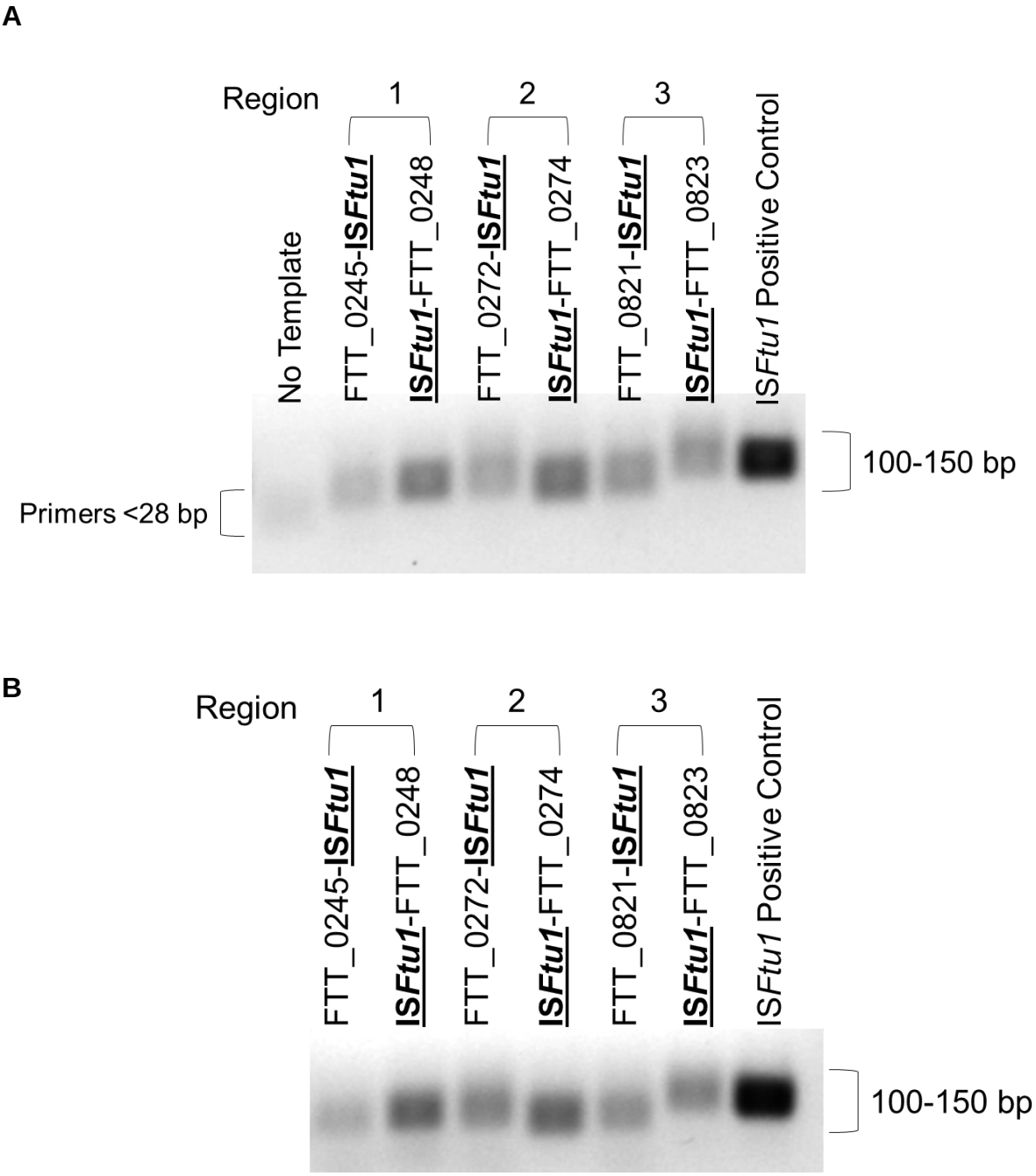


Figure 12: RT-PCR analysis of genes surrounding highly expressed *F. tularensis* ISFtu1 ORFs. Shown are the RT-PCR products obtained for IS*Ftu1* FTT_0247 with either the adjacent 5' and 3' gene transcripts in Region 1 left and right lanes, respectively, FTT_0273 with either the adjacent 5' and 3' gene transcripts in Region 2 left and right lanes, respectively, and FTT_0822 with either the adjacent 5' and 3' gene transcripts in Region 3 left and right lanes for DNase-treated RNA isolated from *F. tularensis* A.I strains **(A)** Schu S4 and **(B)** NE061598, during mid exponential growth phase. *F. tularensis* genomic DNA was used as the template for the positive control, and a no template reaction was used as a negative control.

Next, the relative levels of transcriptional expression for the co-transcribed *ISFtu1* and adjacent genes were compared in *F. tularensis* A.I strains Schu S4 and NE061598, A.II strains WY96 and W4114, attenuated type B LVS, and avirulent subsp. *novicida* strain U112, during mid log growth phase. The adjacent genes encoded by locus tags FTT_0245, FTT_0274, and FTT_0823 were evaluated, since these genes appeared to be co-transcribed with ISFtu1 based on the RNA-Seq data, as well as RT-PCR results (Figure 12). For RNA-Seq assessments, the data was compiled and normalized as previously described (Croucher and Thomson, 2010; Giannoukos et al., 2012).

In all the *F. tularensis* strains examined during mid log growth, transcriptional expression of FTT_0245 (*usp*) was at least 10-fold higher compared to FTT_0274 (hypothetical membrane protein) and FTT_0823 (*mutT*) (Figure 13). However, the virulent A.I and A.II strains expressed less FTT_0245 (*usp*) transcripts than the attenuated type B LVS and avirulent *novicida* strain U112, during mid log growth. The abundance of the FTT_0823 (*mutT*) transcript was higher in the *F. tularensis* subtype A.II strains compared to the subtype A.I strains, type B LVS, and U112. Together these results provide evidence that each of the *F. tularensis* subpopulations exhibit a unique transcriptional profile.

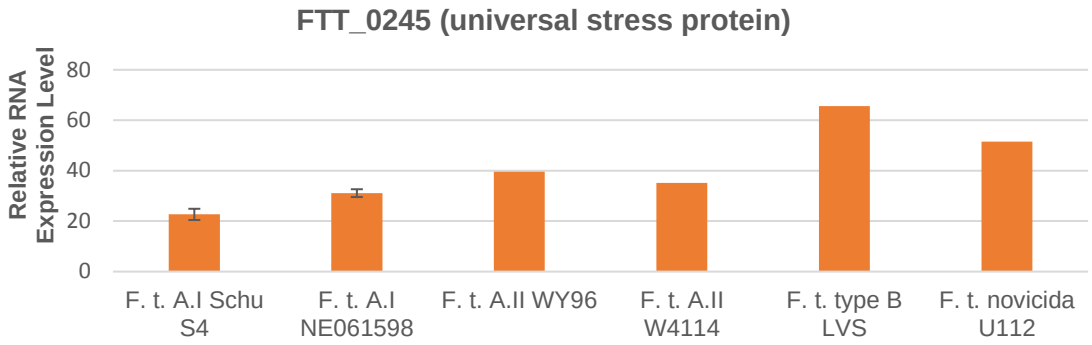
The expression levels of these three genes co-transcribed with IS*Ftu1* were also examined throughout growth in the virulent *F. tularensis* A.I strains Schu S4 and NE061598, as well as attenuated type B LVS and avirulent U112. In general, *usp* (FTT_0245) was expressed at higher levels in all the *F. tularensis* strains examined throughout growth compared to the other two loci examined (Figure 14). The expression levels of *usp* were the highest in attenuated LVS followed by avirulent U112, whereas the type A.I strains produced the lowest abundance of this transcript (Figure 14A). The expression of *usp* was lowest during early log growth phase with slightly higher expression levels during mid and late log growth phase, with the exception of A.I strain NE061598

that had consistent expression throughout growth. FTT_0245 expression levels during stationary phase in Schu S4 and NE061598 remained the same as was observed in mid and late log for the respective strain (data not shown).

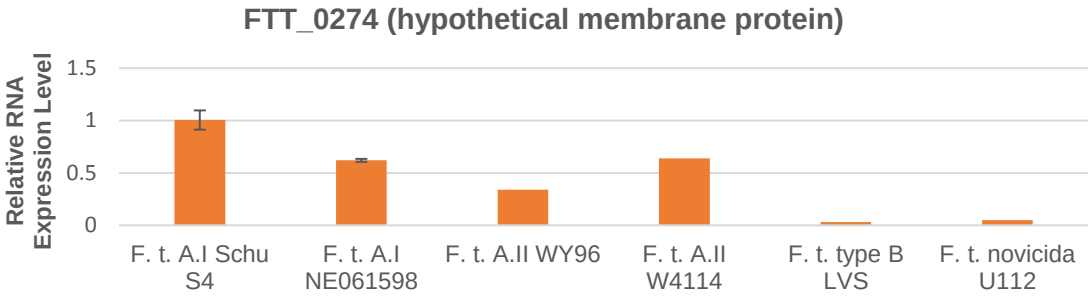
The previously annotated membrane protein FTT_0274 was expressed at the highest level in *F. tularensis* Schu S4 during early and late log growth phase compared to NE061598, LVS, and U112 (Figure 14B). In contrast, there was low or undetectable expression of FTT_0274 throughout growth in NE061598, LVS, and U112. There was no expression of FTT_0274 in NE061598 during stationary phase while Schu S4 had expression levels similar to mid log growth phase during stationary phase (data not shown).

Transcriptional expression of the *mutT* mutator gene (FTT_0823) slightly increased in the *F. tularensis* NE061598, LVS, and U112 strains throughout growth, albeit the abundance remained low (Figure 14C). Schu S4 expression of this gene during growth was inconclusive and will require further study. During stationary phase, the level of *mutT* transcription increased further in NE061598, whereas only a low level of *mutT* expression was observed in Schu S4 (data not shown). Based on these assessments, the co-expressed *usp* and IS*Ftu1* region was selected for further investigation due to (i) the differential expression of *usp* observed in the various *F. tularensis* strains and (ii) the proposed functions of this gene product that allow bacteria to persist in stressful environments.

A



B



C

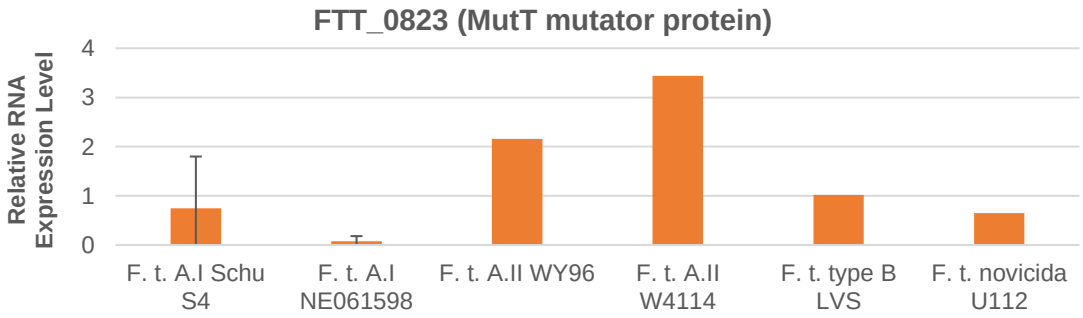
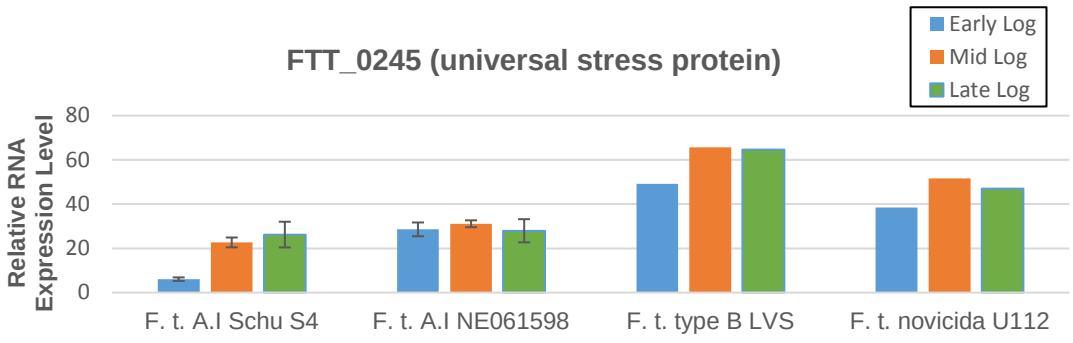
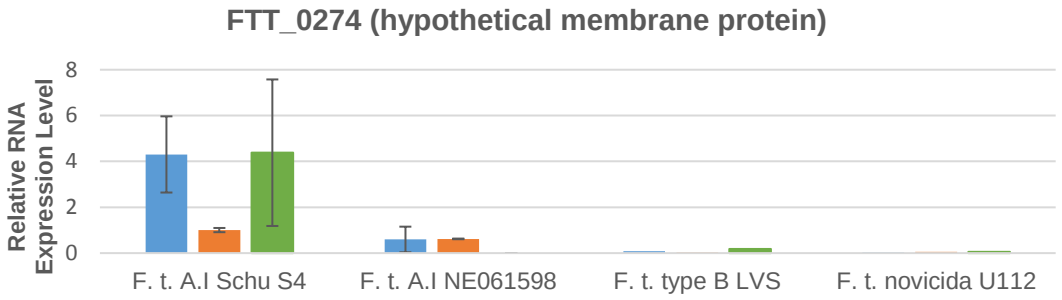


Figure 13: RNA transcript levels for F. tularensis genes co-expressed with IS*Ftu1* during mid log phase. Transcriptional expression of FTT_0245 (universal stress protein), FTT_0274 (hypothetical membrane protein), and FTT_0823 (MutT mutator protein) were examined in *F. tularensis* A.I strains Schu S4 and NE061598, A.II strains WY96 and W4114, attenuated type B LVS, and avirulent subsp. *novicida* U112, during mid log growth phase. RNA-Seq data was used for these analyses. Transcriptomes for *F. tularensis* A.I strains Schu S4 and NE061598 were obtained in two independent experiments and error bars are shown.

A



B



C

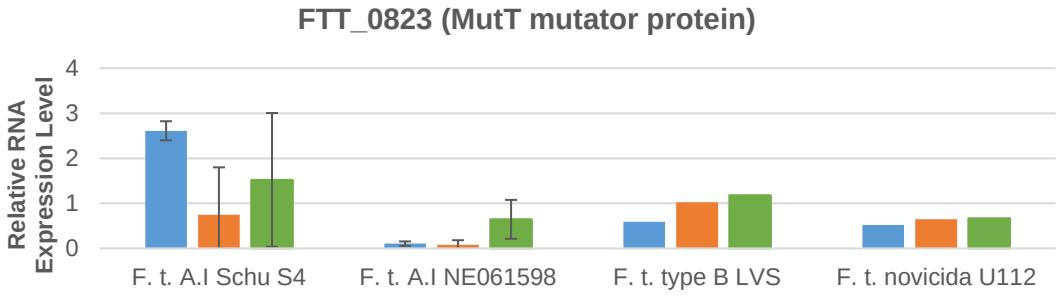


Figure 14: RNA transcript levels for F. tularensis genes co-expressed with IS*Ftu1* during *in vitro* growth. Transcriptional expression of FTT_0245 (universal stress protein), FTT_0274 (hypothetical membrane protein), and FTT_0823 (MutT mutator protein) were examined in *F. tularensis* A.I strains Schu S4 and NE061598, attenuated type B LVS, and avirulent subsp. *novicida* U112, during early, mid, and late log growth phase. RNA-Seq data was used for these analyses. Transcriptomes for *F. tularensis* A.I strains Schu S4 and NE061598 were obtained in two independent experiments and error bars are shown.

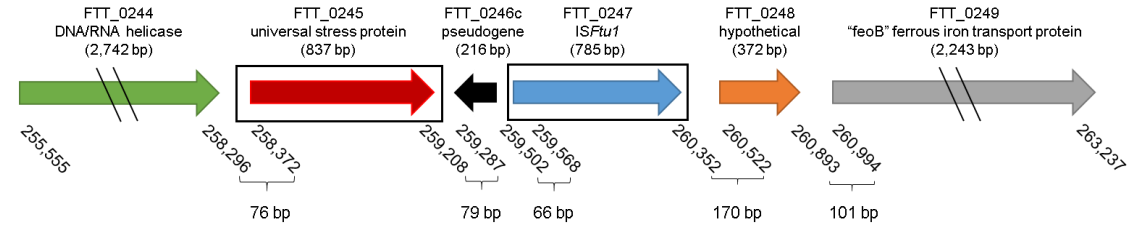
Genetic Analysis of Co-transcribed Universal Stress Protein and IS*Ftu1*

The genetic organization of the genes encoding the universal stress protein and co-expressed IS*Ftu1* was examined in *F. tularensis* A.I strains Schu S4 and NE061598, A.II strains WY96 and W4114, attenuated type B LVS, and subsp. *novicida* U112 (Figure 15). In the A.I strains, the genetic organization of this region was nearly identical, with the only difference being the directionality of a small intervening hypothetical protein or pseudogene. In the A.II strains, this region was identical to each other. However, a comparison between the A.I and A.II genetic organization of the universal stress protein and co-expressed IS*Ftu1* revealed that the directionality, chromosomal location, and flanking genes differed. In attenuated type B LVS, the genetic organization of the universal stress protein and co-expressed IS*Ftu1* gene was more similar to the A.II strains than the A.I strains, but the MFS transporter gene downstream of the IS*Ftu1* appeared to be truncated in LVS. *F. tularensis* subsp. *novicida* U112 is also more similar to the A.II strains than the A.I strains, but lacks the IS*Ftu1* and intervening hypothetical protein genes. These analyses revealed that the context of the co-expressed *usp* and IS*Ftu1* genes differed between the *F. tularensis* A.I, A.II, and B strains, with the A.I strains differing with most with respect to the downstream genes.

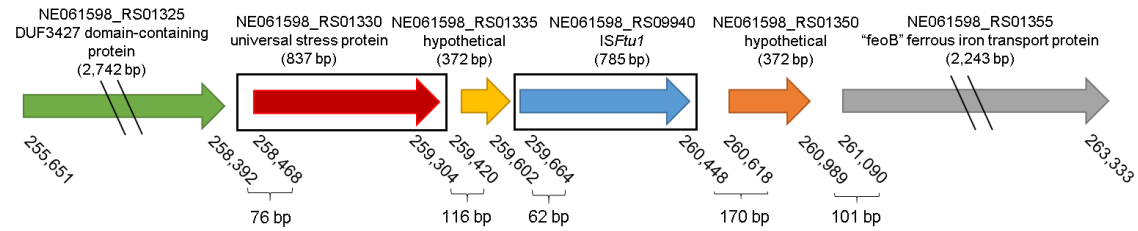
In order to identify the transcriptional start site for the co-expressed *usp* and IS*Ftu1* transcript, 5’ RACE and subsequent sequencing was utilized for the *F. tularensis* A.I strains Schu S4 and NE061598. The results from the sequencing of the 5’ UTR for this transcript are shown in Figure 16 with the adapter sequence and the transcriptional start site denoted. The DNA sequence of the loci with the genes encoding the universal stress protein and IS*Ftu1* in *F. tularensis* Schu S4, as well as the flanking regions are shown in Figure 17. The -35 region, Pribnow box (-10), transcriptional start site (+1), Shine-Dalgarno sequence, translational start codons, and stop codons that were identified are

all indicated in Figure 17. The transcriptional start site that was determined by 5’ RACE was 22 nucleotides downstream of the -10 region and begins 28 bp upstream of the *usp* translational start codon.

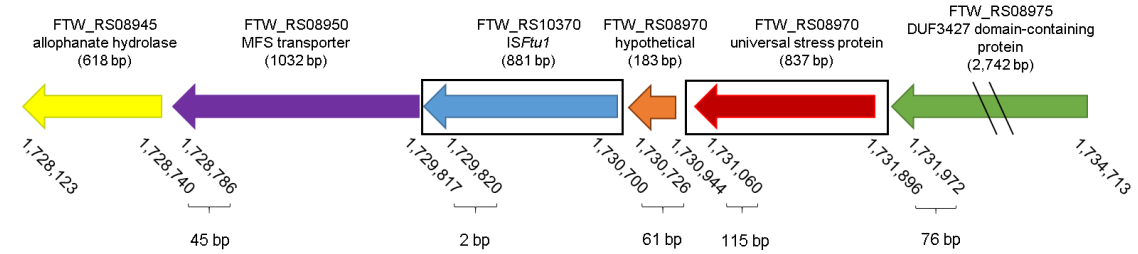
A


FTT_0244
DNA/RNA helicase
(2,742 bp)
FTT_0245
universal stress protein
(837 bp)
FTT_0246c
pseudogene
(216 bp)
FTT_0247
IS*Ftu1*
(785 bp)
FTT_0248
hypothetical
(372 bp)
FTT_0249
"feoB" ferrous iron transport protein
(2,243 bp)
255,555
258,296
258,372
259,208
259,287
259,502
259,568
260,352
260,522
260,893
260,994
263,237
76 bp
79 bp
66 bp
170 bp
101 bp

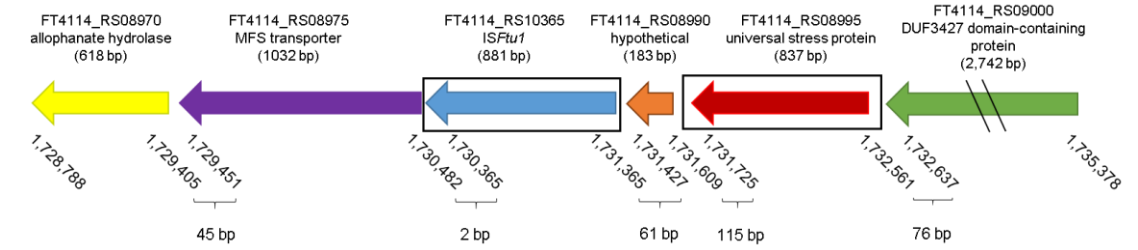
B


NE061598_RS01325
DUF3427 domain-containing protein
(2,742 bp)
NE061598_RS01330
universal stress protein
(837 bp)
NE061598_RS01335
hypothetical
(372 bp)
NE061598_RS09940
IS*Ftu1*
(785 bp)
NE061598_RS01350
hypothetical
(372 bp)
NE061598_RS01355
"feoB" ferrous iron transport protein
(2,243 bp)
255,651
258,392
258,468
259,304
259,420
259,602
259,664
260,448
260,618
260,989
261,090
263,333
76 bp
116 bp
62 bp
170 bp
101 bp

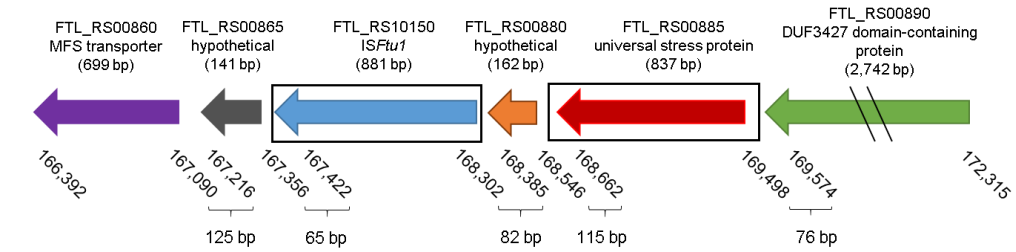
C


FTW_RS08945
allophanate hydrolase
(618 bp)
FTW_RS08950
MFS transporter
(1032 bp)
FTW_RS10370
IS*Ftu1*
(881 bp)
FTW_RS08970
hypothetical
(183 bp)
FTW_RS08970
universal stress protein
(837 bp)
FTW_RS08975
DUF3427 domain-containing protein
(2,742 bp)
1,728,123
1,728,740
1,728,786
1,729,817
1,729,820
1,730,700
1,730,726
1,730,944
1,731,060
1,731,896
1,731,972
1,734,713
45 bp
2 bp
61 bp
115 bp
76 bp

D


FT4114_RS08970
allophanate hydrolase
(618 bp)
FT4114_RS08975
MFS transporter
(1032 bp)
FT4114_RS10365
IS*Ftu1*
(881 bp)
FT4114_RS08990
hypothetical
(183 bp)
FT4114_RS08995
universal stress protein
(837 bp)
FT4114_RS09000
DUF3427 domain-containing protein
(2,742 bp)
1,728,788
1,729,405
1,729,451
1,730,482
1,730,365
1,731,365
1,731,427
1,731,609
1,731,725
1,732,561
1,732,637
1,735,378
45 bp
2 bp
61 bp
115 bp
76 bp

E


FTL_RS00860
MFS transporter
(699 bp)
FTL_RS00865
hypothetical
(141 bp)
FTL_RS10150
IS*Ftu1*
(881 bp)
FTL_RS00880
hypothetical
(162 bp)
FTL_RS00885
universal stress protein
(837 bp)
FTL_RS00890
DUF3427 domain-containing protein
(2,742 bp)
166,392
167,090
167,216
167,356
167,422
168,302
168,385
168,546
168,662
169,498
169,574
172,315
125 bp
65 bp
82 bp
115 bp
76 bp

F

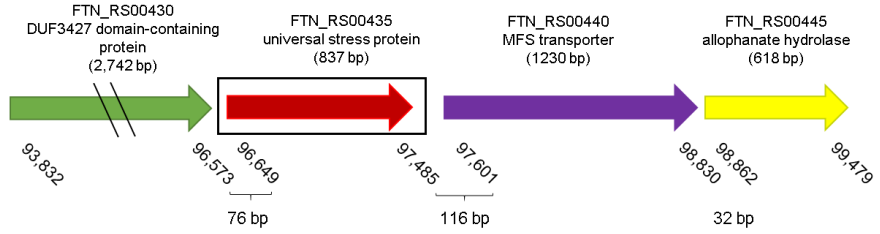

FTN_RS00430
DUF3427 domain-containing protein
(2,742 bp)
FTN_RS00435
universal stress protein
(837 bp)
FTN_RS00440
MFS transporter
(1230 bp)
FTN_RS00445
allophanate hydrolase
(618 bp)
93,832
96,573
96,649
97,485
97,601
98,830
98,862
99,479
76 bp
116 bp
32 bp

Figure 15: Context of the co-expressed gene encoding a universal stress protein and IS*Ftu1* in *F. tularensis* strains. Diagram of highly co-expressed universal stress protein (FTT_0245, red boxed arrow) and IS*Ftu1* (FTT_0247, blue boxed arrow), along with surrounding genes in *F. tularensis* **(A)** A.I strain Schu S4, **(B)** A.I NE061598, **(C)** A.II WY96, **(D)** A.II W4114, **(E)** attenuated type B LVS, and **(F)** subsp. *novicida* U112. The directionality of gene expression is indicated with an arrow. The beginning and ending nucleotide location of each gene in the respective *F. tularensis* chromosome is denoted underneath, along with the distance between each gene.

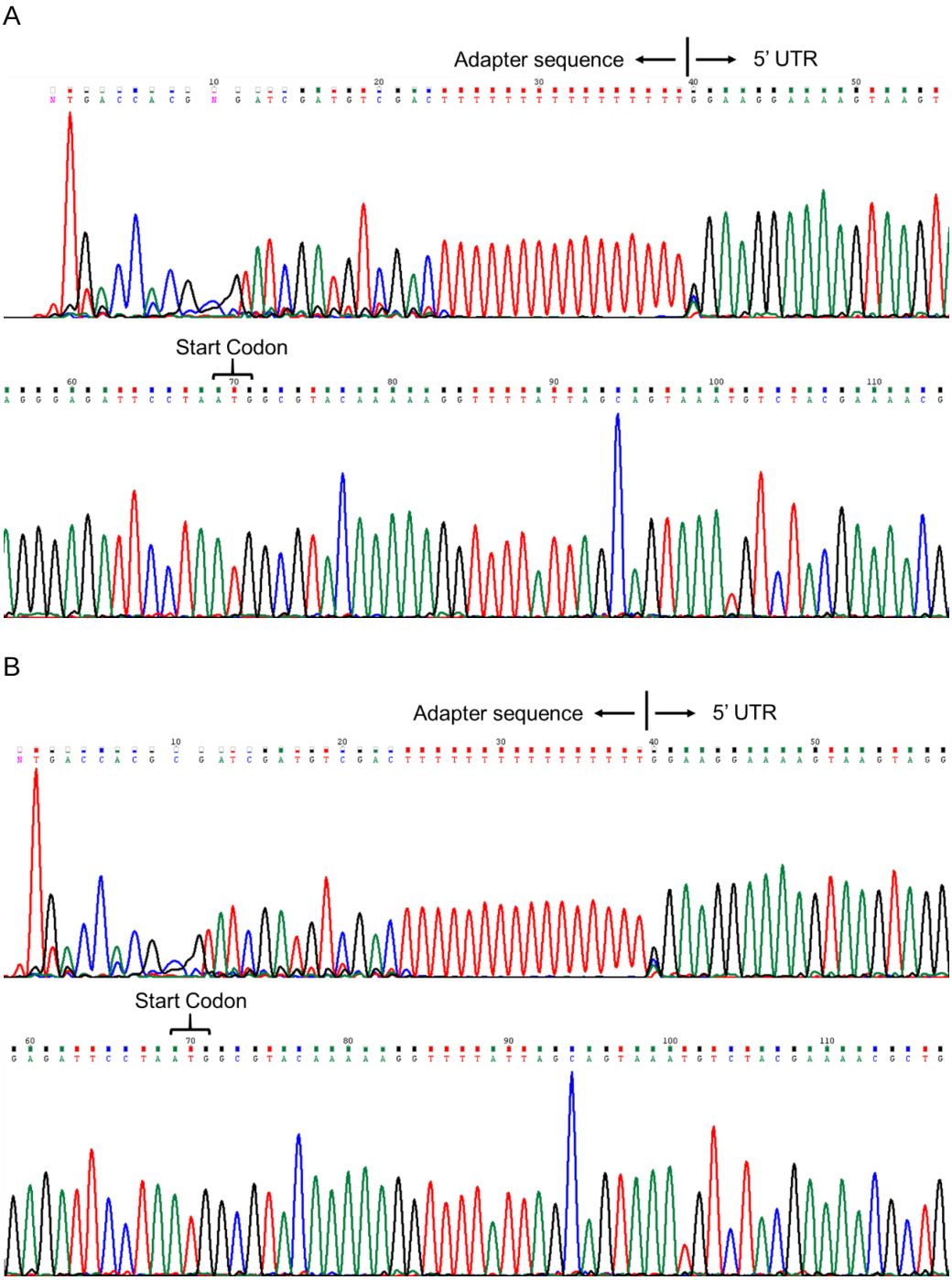


Figure 16: 5' RACE sequence for the co-expressed *F. tularensis* universal stress protein gene and downstream IS*Ftu1*. (A) Chromatogram showing the 5' RACE sequence for the universal stress protein (FTT_0245) and IS*Ftu1* (FTT_0247) co-transcribed genes in *F. tularensis* A.I strains **(A)** Schu S4 and **(B)** NE061598. The adapter sequence and 5' untranslated region are indicated with arrows above the respective sequence. The translational start codon for the universal stress protein gene is also denoted above the associated sequence.

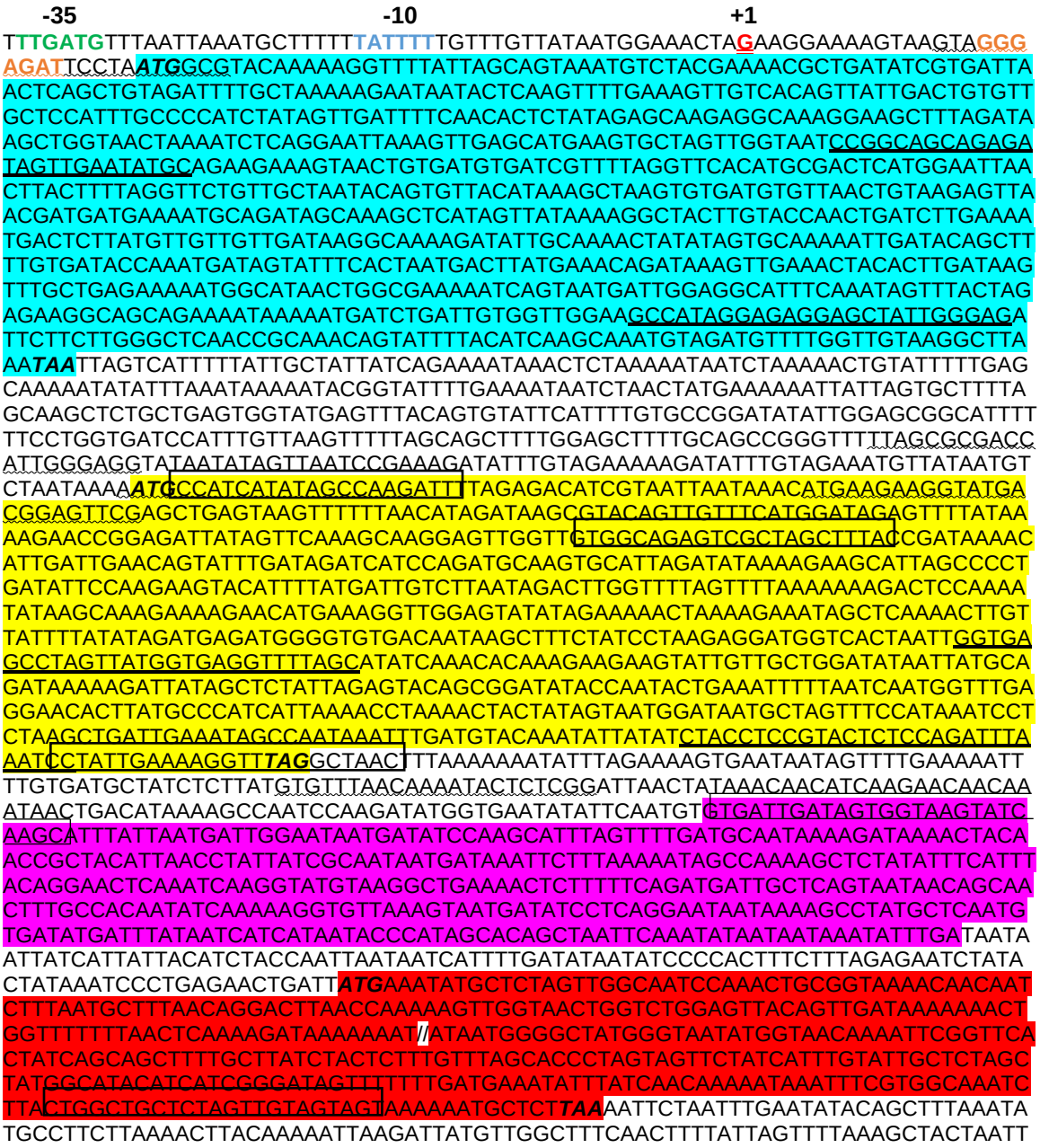


Figure 17: Nucleotide sequence of the universal stress protein gene and IS*Ftu1*, including flanking regions. The *F. tularensis* Schu S4 bicistronic transcript encoding the FTT_0245 universal stress protein (blue), FTT_0247 IS*Ftu1* gene (yellow), FTT_0248 hypothetical protein (purple), and FTT_0249 ferrous iron transport (*feoB*) protein (red) are shown. The region in between FTT_0245 and FTT_0247 has a possible hypothetical protein or a hypothetical pseudogene (FTT_0246). The transcriptional start site (+1) is double underlined and in red bold text. The start and stop codons for each gene is in bold and italicized. The predicted -35 region (green text), Pribnow box (-10, blue text), and Shine-Dalgarno sequences (orange text) are also denoted in the 5' UTR. Forward primers (wavy underline) and reverse primers (boxed) used to generate the results for Figures 12, 18, and 19 are marked. Primers used for 5' RACE are in bold and underlined.

To further confirm the co-expression of the genes encoding the universal stress protein (FTT_0245) and IS*Ftu1* (FTT_0247), as well as to determine if additional downstream genes were included in this transcript, RT-PCR was performed using appropriate upstream and downstream primers. The expected 2 kb RT-PCR product was obtained with the *usp* and IS*Ftu1* specific primers, verifying the co-transcription of these two genes (Figure 18). To determine if additional downstream genes were also co-expressed with *usp* and IS*Ftu1*, primer pairs that amplify IS*Ftu1* or *usp* and a region in the 3' end of the adjacent downstream *feoB* ORF (FTT_0249) were evaluated in RT-PCR. In these preliminary assessments, no RT-PCR products were obtained (Figure 18), indicating that only a bicistronic transcript was produced for *usp* and IS*Ftu1*.

Co-transcription of the complete genes encoding *usp* (FTT_0245) and IS*Ftu1* (FTT_0247) throughout growth for *F. tularensis* A.I strains Schu S4 and NE061598 was next evaluated using RT-PCR. In these assessments, the expected 2 kb RT-PCR product for the bicistronic *usp* and IS*Ftu1* transcript was detected in early and mid log growth phase, but surprisingly not during late log or stationary phase for both A.I strains (Figure 19). However, RNA-Seq data indicated that both *usp* and IS*Ftu1* were still being expressed at moderate to high levels during late log and stationary phase. Collectively, these results indicated that the co-expression of *usp* and IS*Ftu1* was induced during early growth *in vitro* when nutrients are not limited, but interestingly during late log growth and stationary phase when medium components become depleted these two genes were transcribed separately. Therefore, transcriptional termination occurs after *usp* and prior to the co-transcription with IS*Ftu1*, during late log growth and stationary phase by an unknown mechanism. The intervening 359 bp sequence between *usp* and IS*Ftu1* does not contain any obvious stem loop structure for intrinsic transcriptional or rho independent

termination, but does have a lower G+C content (28%) than the overall *F. tularensis* NE061598 genome (32.3%, Table 3).

To evaluate if nitric oxide and polyamine effected the co-transcription of *usp* and IS*Ftu1*, as well as the expression of other genes in *F. tularensis* A.I NE061598, the bacterial cells were treated with SNAP or spermine during early log growth phase. As shown in Figure 19, RT-PCR showed that the A.I strains treated with SNAP or spermine produced approximately 2-fold higher levels of the *usp*/IS*Ftu1* bicistronic transcript compare to untreated bacterial cells. Therefore, both nitric oxide and the polyamine spermine induce higher co-expression levels of *usp* and IS*Ftu1* in the wild type *F. tularensis* A.I strain examined during early growth phase.

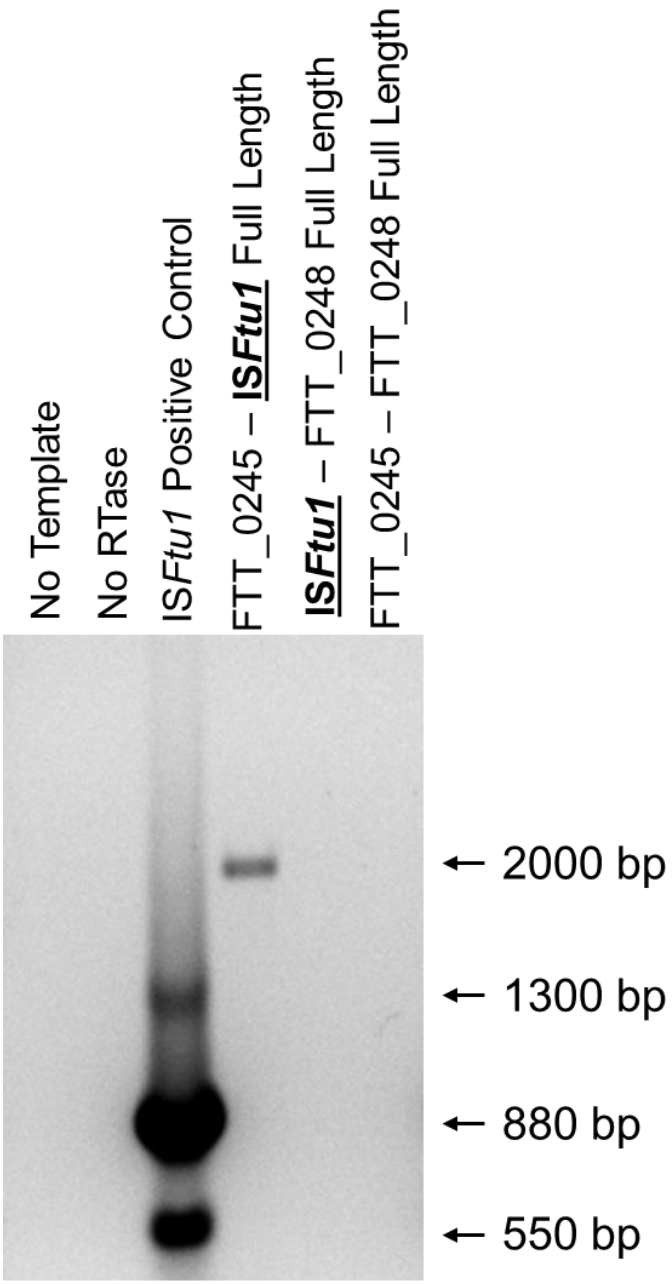


Figure 18: RT-PCR analysis of highly expressed transcripts encoding the universal stress protein (Usp) and downstream IS*Ftu1* in *F. tularensis* A.I strain Schu S4. RT-PCR of *usp* (FTT_0245) and IS*Ftu1* (FTT_0247), IS*Ftu1* and FTT_0248, and both adjacent genes FTT_0245 and FTT_0248 to detect co-transcribed products. *F. tularensis* genomic DNA was used as the template in the positive control, and negative controls included a reaction without template and a reaction without RTase.

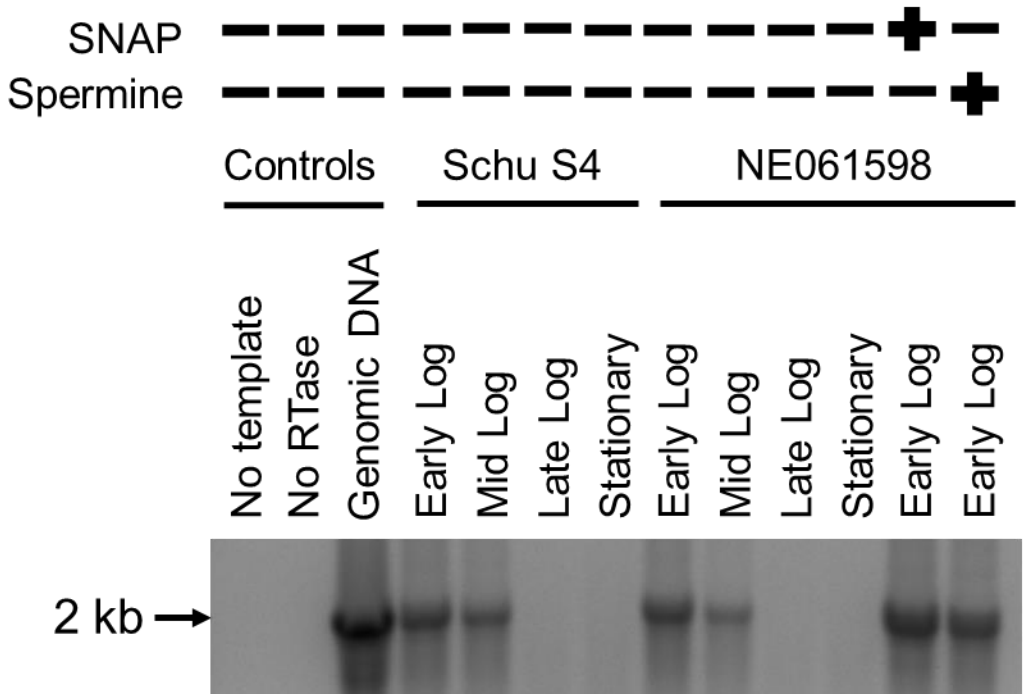
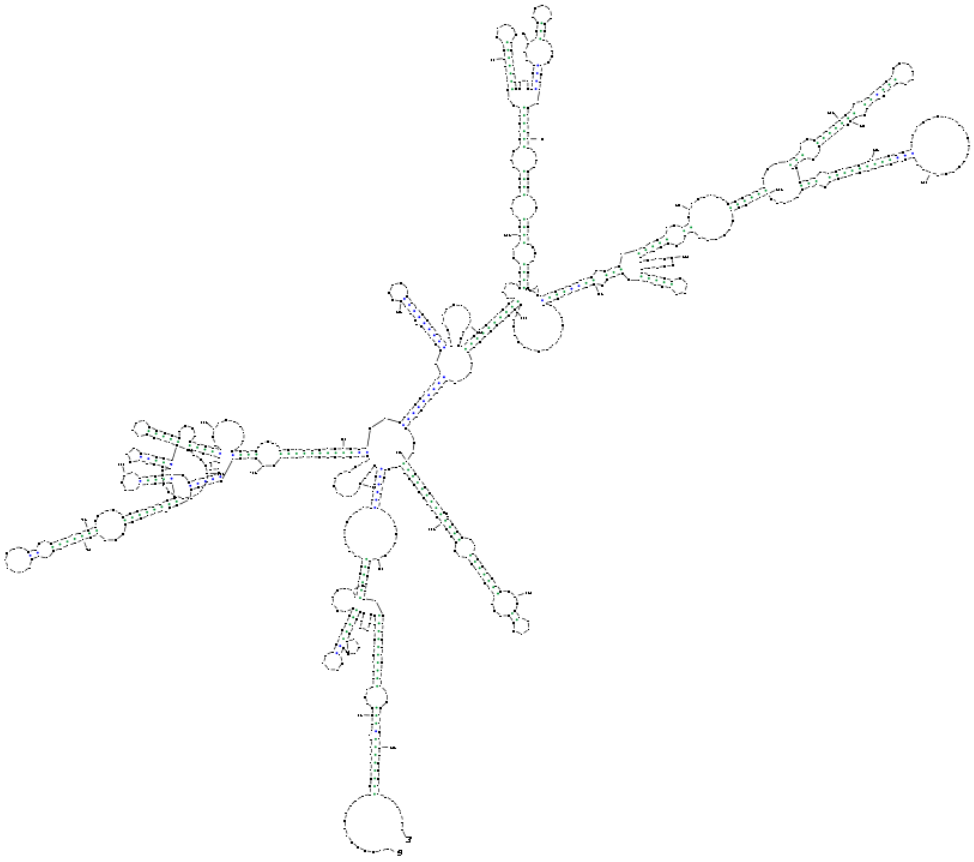


Figure 19: RT-PCR analysis of universal stress protein gene and co-expressed IS*Ftu1* throughout growth for *F. tularensis* A.I strains Schu S4 and NE061598 without and with SNAP or spermine. The growth phase in which total RNA was isolated from the *F. tularensis* A.I strains is shown above each associated lane. The *usp* specific (FTT_0245) forward and IS*Ftu1* (FTT_0247) reverse primers are shown in Table 2 and the location is marked in Figure 17. *F. tularensis* genomic DNA was used as the template in the positive control, and negative controls included a no template reaction and a no RTase reactions.

The predicted secondary structure of the transcript encoding the universal stress protein alone or co-expressed with IS*Ftu1* was next assessed using the software Sfold 2.2. In theory, these minimum free energy models determine the secondary RNA structure that will most likely form in nature, and the structure that utilizes the least amount of energy will be a more favorable process. Accordingly, these models indicated that the bicistronic transcript encoding *usp* and IS*Ftu1* will form a higher-order RNA structure much more readily with a ΔG of -458 (Figure 20) compared to the monocistronic *usp* transcript, which had a ΔG of -204 (Figure 21). Additional secondary RNA structure modeling softwares such as RNAfold and RNAstructure were used and provided similar results.



ΔG°_{37} = -204.20

Figure 20: Predicted minimum free energy structure of the transcript encoding only the universal stress protein without IS*Ftu1* co-expression. The structure of the monocistronic mRNA encoding the universal stress protein shown was predicted by the Sfold 2.2 software.

Figure 21: Predicted minimum free energy structure of the transcript encoding the universal stress protein that is co-transcribed with IS*Ftu1*. The bicistronic structure of the mRNA encoding the universal stress protein and IS*Ftu1* shown was predicted by the Sfold 2.2 software.

DISCUSSION

F. tularensis Genotypic and Phenotypic Comparisons

Genome size and % G+C content, as well as protein and gene totals, are very similar for the various *F. tularensis* clades and strains (see Table 3). Interestingly, *F. tularensis* subsp. *novicida* U112 contains more predicted proteins than the other strains, despite having a lower number of predicted genes, with the exception of Schu S4. The opposite case was observed for LVS, with the highest number of predicted genes and a lower number of predicted proteins, with the exception of Schu S4. NE061598, WY96, and W4114 had a similar number of predicted genes to predicted proteins. The largest genotypic difference between the subspecies is with the number of pseudogenes. *F. tularensis* subsp. *novicida* contains only 49 pseudogenes, while the type A and type B strains contain over 200 pseudogenes. This high number of pseudogenes may be indicative of genome reduction in the virulent subspecies as they adapted to an intracellular life style (Casadevall, 2008).

F. tularensis attenuated LVS and avirulent U112 grew approximately twice as fast as the virulent type A.I and type A.II strains, doubling every 1.4 hours in supplemented BHI (see Figure 2). *F. tularensis* U112 belongs to the subsp. *novicida* with more functional proteins and less pseudogenes than the other subspecies, and therefore would be expected to grow at a faster rate. *F. tularensis* LVS is an attenuated strain that was produced from a fully virulent type B strain by the Russians, using classical bacteriological methods (i.e., selection of individual bacterial colonies and repeated subcultures under conditions stressful to the bacterium) (Pollitzer, 1967). These laboratory procedures apparently led to mutations in the LVS genome that were both incidental and responsible for virulence attenuation (Rohmer et al., 2006), and may have contributed to the selection of genes that promoted faster growth *in vitro*. In contrast, the virulent *F. tularensis* strains are known to be fastidious and grow slow, doubling approximately every 3-4

hours. The slower *in vitro* growth of the virulent strains may reflect an adaptation to an intracellular lifestyle, specifically the loss of metabolic pathway genes (Casadevall, 2008).

F. tularensis Type A.I Response to Nitric Oxide

The ability of a pathogen to appropriately respond to environmental cues is essential for its survival. Macrophages produce nitric oxide during an infection (MacMicking et al., 1997), and are the primary cells infected by *F. tularensis* (Carvalho et al., 2014). Reactive nitrogen is critical for the clearance of *F. tularensis* LVS, but is insufficient to control *F. tularensis* Schu S4 intracellular survival and replication (Anthony et al., 1992; Newstead et al., 2014). Nitric oxide can react with superoxide and produce peroxynitrite ($ONOO^-$), which is important for microbicidal killing of pathogens (Darrah et al., 2000). Nitric oxide produced by SNAP has also been shown to contain *F. tularensis* LVS in the macrophage phagosome and substantially inhibit *mglA* expression, but did not decrease this pathogen's viability (Tancred et al., 2011). In another study, *F. tularensis* Schu S4 suppressed nitric oxide production by two different types of macrophages, unlike LVS (Newstead et al., 2014); however, the mechanisms by which this pathogen suppresses reactive nitrogen species and evades an effective immune response are unknown.

In the current study, we evaluated the response of a wild type *F. tularensis* A.I strain to SNAP, a nitric oxide donor. The addition of SNAP to wild type A.I NE061598 during early log growth phase did not have any observable effect on growth (Figure 3), but did importantly upregulate IS*Ftu1* expression (Table 12). However, the co-expressed genes with IS*Ftu1* that this investigation focused on, specifically *usp* (FTT_0245), a hypothetical membrane protein (FTT_0274), and *mutT* (FTT_0823) did not show any obvious changes in expression levels in response to nitric oxide. Other genes that were upregulated in response to nitric oxide included seven genes encoding lipoproteins (data not shown). Lipoproteins are membrane proteins that can be involved in various functions, such as translocation of virulence factors, surface adhesion,

antigenic variation, and the induction of the host inflammatory response (Kovacs-Simon et al., 2011). Another gene of interest that was upregulated in response to SNAP was the capsule biosynthesis protein C gene (*capC*). The capsule is important to protect *F. tularensis* from the host detection pathways, and a study by others showed that increased capsule production in LVS led to a more virulent strain (Bandara et al., 2011; Lindemann et al., 2011). Although SNAP was shown to substantially (>50-fold) inhibit *mglA* expression in attenuated type B LVS (Tancred et al., 2011), our results did not show a significant (<2-fold) difference in *mglA* transcript levels in wild type A.I strain NE061598 with the addition of SNAP.

F. tularensis Type A.I Response to Polyamine

Spermine is present at millimolar levels in the cytosol of eukaryotic cells and is not produced by bacteria (Igarashi and Kashiwagi, 2000). Spermine is a polyamine that can stabilize secondary structures, as well as regulate gene expression and promote DNA, RNA, and protein synthesis (Yoshida et al., 2004). Carlson and associates demonstrated that spermine can upregulate IS elements in *F. tularensis* LVS and Schu S4 (Carlson et al., 2009). In these studies, microarray analyses and IS promoter reporter constructs were used to study *F. tularensis* gene expression in the murine macrophage cell line RAW 264.7. This investigation also determined that the proinflammatory response by macrophages was reduced when infected with *F. tularensis* that was previously grown in medium containing spermine compared to medium lacking spermine.

To address this outcome, a transposon screen in attenuated LVS was implemented to identify mutants that were unable to respond to extracellular spermine (Russo et al., 2011). Their results showed that FTL_0883 in LVS and the homolog in Schu S4 (FTT_0615c), which is predicted to encode a metal ion transporter, are induced by spermine and contribute to *F. tularensis* virulence by sensing the host environment. More specifically, in-frame deletions of this gene in LVS and Schu S4, along with complementation studies, demonstrated an important role

for the respective gene product in suppressing the production of proinflammatory cytokines by infected macrophages. However, although FTL_0883 was needed for wild type rates of LVS replication within macrophages, this gene in Schu S4 was not needed for replication within macrophages. Therefore, the function of the protein in Schu S4 remains undefined, emphasizing the phenotypic differences between virulent and attenuated *F. tularensis* strains.

Investigations by others have demonstrated that host polyamine biosynthesis pathways can be activated by pathogens and contribute to the suppression of apoptosis in infected macrophages, sustaining the pathogen's replication niche (Andersson et al., 2006; Bergman et al., 2005). For example, a lung infection in mice by a *F. tularensis* type A strain increased host synthesis of ornithine decarboxylase, a key and early step in polyamine synthesis that contributes to spermine accumulation (Andersson et al., 2006). However, additional study is needed to understand the effect of host polyamines on *F. tularensis* pathogenicity.

To assess the effect of spermine on the expression of IS*Ftu1* and other genes in a wild type A.I strain, we added spermine to the supplemented BHI medium, during early log growth phase of *F. tularensis* NE061598. Although there was not a notable effect on growth rate, with doubling occurring approximately every 3.8 hours (Figure 3), there were noteworthy differences in gene expression for the untreated versus spermine treated cultures. RNA-Seq was used to examine the effects of spermine on gene expression, which is a more comprehensive approach than the previous microarray work done by Carlson and associates (Carlson et al., 2009).

Our data showed that spermine upregulated IS*Ftu1* expression levels in *F. tularensis* NE061598 (Table 12), corroborating with the findings by Carlson et al. (2009). In *F. tularensis* NE061598, the *Francisella* spermine response genes *fsr1* (Schu S4 FTT_1480c homolog), *fsr2* (Schu S4 FTT_1163c homolog), and *fsr3* (Schu S4 FTT_0018 homolog) were upregulated and the spermidine synthase gene *speE* (Schu S4 FTT_0431 homolog) was downregulated when spermine was added to the cultures, again in agreement with the data obtained by Carlson and

associates in LVS and Schu S4. However, since the study by Carlson et al. predominantly focused on LVS, many of the spermine responsive genes that were located by an IS element were not relevant in NE061598, due to the numerous differences in genomic organization between the *F. tularensis* clades.

Interestingly, our results showed that the *mutT* gene (FTT_0823), which was co-expressed with IS*Ftu1* in *F. tularensis* NE061598, was downregulated in response to spermine, although the levels of expression were low without or with spermine treatment. No change in expression levels were observed in response to spermine for the other genes co-expressed with IS*Ftu1* that were examined in more detail, which included the genes encoding a universal stress protein (FTT_0245) or a hypothetical membrane protein (FTT_0274).

Previous work by Russo et. al. determined that FTT_0615c, a metal ion transporter, was upregulated in response to spermine in the *F. tularensis* A.I prototype strain, Schu S4 (Russo et al., 2011). In contrast, our results showed a 2-fold down regulation of this homolog in the wild-type A.I strain NE061598 with the addition of spermine. These differences could be due to differences in the culturing of these select agent strains and/or inherent differences in the laboratory strain Schu S4 versus the wild-type strain NE061598, since the number of passages and source of these pathogens is important.

Phagocytes, such as macrophages, lymphocytes, and natural killer cells all require polyamines, and polyamines, like spermine, play a role in differentiation and the synthesis of proteins. (Seiler, 1999). Spermine production by eukaryotic cells has been shown to play a role in DNA synthesis, transcription, and translation, as well as reduce the cytokine response from infected macrophages (Carlson et al., 2009). The findings presented in this study and those of others provided evidence that spermine indeed upregulates *F. tularensis* IS element expression with or without the adjacent gene. Although there is increasing evidence that polyamines influence bacterial virulence in host-pathogen interactions, the mechanisms by which all of these processes

occur are unknown and are undoubtedly complicated by multiple environmental signals from both *F. tularensis* and the infected host.

F. tularensis Insertion Sequence Elements

The different *F. tularensis* subspecies, subtypes, and strains share many features with regard to genome size and average nucleotide identity (ANI). However, the subspecies and subtypes differ considerably in virulence. For example, although the *F. tularensis* A.I and A.II strains share the highest ANI at 99.6%, our previous study revealed important genotypic differences between these two subtypes that may contribute to disparities in virulence (Larson et al., 2015). Therefore, we propose that many of the pathogenicity differences observed between the *F. tularensis* subpopulations are due to chromosomal translocations, which occurred predominately between the numerous IS elements, and that these modifications have contributed to altered gene regulation and expression. Although this study examined the upregulated ISFtu1 ORF and adjacent genes, the downregulated or disrupted genes due to an IS element undoubtedly also contributes to the differential expression within the *F. tularensis* clades.

Many highly virulent pathogens contain numerous IS elements, including *F. tularensis*. IS elements can contribute to genome reduction and rearrangements, and alter or disrupt the expression of nearby genes. The most abundant IS element in the virulent subspecies of *F. tularensis* is IS*Ftu1*. This mobile element has undoubtedly contributed to the apparent chromosomal rearrangements and pathogenicity differences noted within this species for the various subpopulations. However, the consequences of the numerous IS*Ftu1* in this pathogen are ill-defined. This study provided insight into the functional role of IS*Ftu1* on *F. tularensis* gene expression.

The most abundant IS element in virulent *F. tularensis* clades is IS*Ftu1*, which has several auto-regulation mechanisms to prevent over-expression and potential destabilization of the

genome. The presences of a -1 programmed translation frameshift required to encode the entire transposase is one such mechanism (Figure 4). If the ribosome fails to slip 1 nucleotide backwards in the slippery heptamer region, then the second half of the protein that contains the DDE active site will not be produced, rendering the transposase inactive (Lohe et al., 1997). The A.I strains have a premature stop codon in all of the IS*Ftu1* open reading frames, even though the DDE motif would still be present if the -1 programmed translational frameshift occurred. In addition, the direct repeat in the middle of IS*Ftu1* gene could also play a role in the inactivation of IS*Ftu1*, by mediating a 312 bp deletion that encompasses both aspartic acid residues of the catalytic DDE motif (Figure 6). This deletion is found in several of the IS*Ftu1* genes within the genomes of both the virulent type A.I and A.II strains, but not in the type B strains including attenuated LVS. These analyses collectively indicate that the IS*Ftu1* open reading frames within the *F. tularensis* type A strains are being inactivated, contributing to genome stabilization for this highly virulent subpopulation.

Examination of the predicted protein structure for the IS*Ftu1* transposase from A.I Schu S4, A.II WY96, and type B LVS was performed. All three IS*Ftu1* protein models aligned the closest to Mariner Mos1 transposase with roughly 15% sequence identity, a member of the IS*630* Tc-1 mariner family (Mahillon and Chandler, 1998). The previously described amino acid differences did not appear to have any overtly obvious effect on the predicted protein structure of IS*Ftu1* between these three strains; however, whether these differences in residue composition affect IS*Ftu1* transposase activity is unknown and requires further study.

IS*Ftu1* Transcript Analysis

Our results demonstrated that the IS*Ftu1* transcript is being synthesized in the *F. tularensis* A.I and A.II strains. Higher expression levels of the 5’ region compared to the 3’ region of IS*Ftu1* was observed in the analysis of the RNA-Seq data (Table 12); however, qRT-PCR is

need to confirm these results. The full-length IS*Ftu1* transcript was detected in the northern blot analysis of the prototype A.I strain Schu S4, but not in wild type A.I strain NE061598, probably due to low sensitivity of DIG-labeled versus ^{32}P-labeled probes (Streit et al., 200p). Nevertheless, the full-length IS*Ftu1* transcript in the both A.I strains Schu S4 and NE061598 was observed when RT-PCR was performed and these results demonstrated that the expression of this gene throughout the entire growth cycle (Figure 8B). In addition to the full-length IS*Ftu1* transcript, a truncated RNA that was 570 bp in length was being produced, most likely representing the IS*Ftu1* RNA with the 312 bp deletion. Since these results determined that constitutive expression of the full-length IS*Ftu1* transcript was occurring throughout growth, we next sought to determine if antisense IS*Ftu1* RNA was also being produced, in order to prevent genome instability. Antisense is one mechanism of auto-regulation that represses expression of transposases during normal conditions (Thomason and Storz, 2010). Our findings indicated that IS*Ftu1* antisense RNA is being synthesized in both *F. tularensis* Schu S4 and NE061598 throughout the growth cycle (Figure 8C), indicating a strategy by which this pathogen prevents deleterious chromosomal rearrangements. Since transcriptional expression of IS*Ftu1* occurs during a macrophage infection (Carlson et al., 2009), and our data shows the constitutive transcription of IS*Ftu1* throughout growth, this expenditure of energy may provide a fitness advantage by regulating the expression or stability of other transcripts, similar to the expression of *usp* and the downstream IS*Ftu1*.

Co-expression of IS*Ftu1* and Adjacent Genes

To identify potential co-expressed genes with an IS*Ftu1*, RNA-Seq results that showed expression in both the IS*Ftu1* ORF and adjacent gene were assessed, due to the high copy number and nucleotide identify of this IS element in the virulent *F. tularensis* strains. In *F. tularensis* A.I strains Schu S4 and NE061598, there were a total of 50 out of 50 and 45 out of 47 respectively, full length or partial IS*Ftu1* ORFs expressed. There were 13 moderately or highly expressed IS*Ftu1* genes in Schu S4, with 7 (53.8%) of those IS elements containing 5’ and/or 3’

adjacent genes with moderate or high expression (Table 4). In contrast, 29.8% of the adjacent genes showed moderate or high expression in *F. tularensis* NE061598 and there was only one moderately expressed IS*Ftu1* gene; however, both the 5' and 3' adjacent genes were highly expressed at this location (Table 5). These highly expressed adjacent genes were a molecular chaperone high-temperature protein G (HtpG) and a hypothetical protein. HtpG is a chaperone that helps proteins fold properly under stressful conditions (Buchner, 2010). HtpG interacts with DnaK/DnaJ/GrpE complex, as well as with DnaA (Grudniak et al., 2015). This IS*Ftu1* ORF in *F. tularensis* Schu S4 had low expression in the 5' region and no expression in the 3' region. However the *htpG* gene was highly expressed in Schu S4, while the hypothetical protein had no expression. These results demonstrated that the A.I prototype strain Schu S4 and the A.I wild type NE061598 differ in the expression levels of ISFtu1 and the adjacent genes.

F. tularensis A.II strain WY96 contained 20 moderately or highly expressed IS*Ftu1* genes, with 15 (42.5%) of those IS elements containing 5' and/or 3' adjacent genes with moderate or high expression (Table 6). *F. tularensis* A.II strain W4114 had 32 moderately or highly expressed IS*Ftu1* ORFs, with 22 (43.8%) of the adjacent genes having moderately or highly expressed 5' and/or 3' adjacent genes (Table 7). LVS had two moderately expressed IS*Ftu1* genes (Table 8). The 5' adjacent gene of the first moderately expressed ISFtu1 encoded a fumerate hydratase (FTL_RS02755) and had high expression, while the 3' adjacent gene had no expression. In the second ISFtu1 ORF, there was no expression of the 5' adjacent gene and moderate expression of the 3' adjacent gene, which encoded an IS5/IS1182 family transposase (FTL_RS10060). The only copy of IS*Ftu1* in subsp. *novicida* strain U112 was expressed at a low level and there was no expression by the adjacent genes (Table 9). Although the A.II strains had similar expression levels of IS*Ftu1* genes and the adjacent genes, the attenuated LVS and avirulent U112 strains differed considerably compared to each other and the virulent *F. tularensis* strains.

F. tularensis NE061598 treated with SNAP contained 20 moderately expressed IS*Ftu1* genes, with 16 (37.5%} of those ORFs containing 5' and/or 3' adjacent genes with moderate or high expression (Table 10). When NE061598 was grown in the presence of spermine, there were 13 moderately or highly expressed IS*Ftu1* genes, with 9 (46.2%) of these ORFs containing 5' and/or 3' adjacent genes with moderate of high expression (Table 11). Both SNAP and spermine increased the overall amount of IS*Ftu1* expression, as well as the expression of the 5' and 3' adjacent genes in NE061598. However, the contribution of these upregulated genes on *F. tularensis* physiology and virulence remains to be determined.

Three genes potentially co-transcribed with IS*Ftu1* were initially selected for additional evaluation in this study. These genes encoded a universal stress protein (Usp), a hypothetical membrane protein, and a MutT protein. The universal stress protein is a member of the conserved UspA family and is induced by exposure to antibiotics, nutrient starvation, heat shock, and other stressful conditions (Nystrom and Neidhardt, 1992; Siegele, 2005). UspA has also been shown to protect bacteria from DNA damage and superoxide stress, as well as involved with adhesion in some bacterial species (Siegele, 2005). The hypothetical membrane protein was of interest since these proteins could be involved with adhesion, transporting molecules across the membrane, signal receptors, or other functions (Nam et al., 2009), and are potential drug targets (Overington et al., 2006). The MutT protein was selected to evaluate since MutT helps to ensure accurate replication and transcription within bacteria. MutT specifically prevents A-(8-oxoG) mispairings from occurring that are not normally repaired through mismatch repair or proofreading (Fowler and Schaaper, 1997). However, since the *usp* gene was confirmed to be co-expressed with IS*Ftu1*, was differentially expressed between the various *F. tularensis* subpopulations, and has been attributed to numerous functions during stress, this locus was selected for a more detailed assessment.

All strains examined contained the conserved gene encoding the universal stress protein, followed downstream by a small hypothetical protein, and then an IS*Ftu1* gene, with the exception of *novicida* U112 (Figure 15). The gene upstream of *usp*, a DUF3427 domain-containing protein, is also the same in all strains, including U112. However, the genes downstream from IS*Ftu1* vary depending on the subspecies. The genetic organization for this locus was shared between the same subtype, but the orientation of the genes differed and may be a consequence of the IS*Ftu1* presence at this location and the abundance of this IS element in the chromosome of virulent *F. tularensis* strains.

Transcriptional Expression of Bicistronic IS*Ftu1* and the Universal Stress Protein

Our study demonstrated that the *usp* gene and the downstream IS*Ftu1* gene are being co-transcribed in the *F. tularensis* A.I strains during early and mid log growth phase, but not in late log or stationary phase (Figure 18 and 19).). However, there was transcriptional expression of both *usp* and IS*Ftu1* throughout the growth cycle of *F. tularensis*, including late log growth and stationary phase (Figure 14 and data not shown). Therefore, the co-transcription of *ups* with IS*Ftu1* may protect the *usp* transcript from degradation by the inherent structure and/or the binding of protective proteins. IS*Ftu1* may also be sequestering *usp* from the translation machinery until needed when nutrients are becoming depleted. Once the bacteria reaches late log growth and stationary phase, the co-expression of *ups* along with IS*Ftu1* may no longer be an advantageous to the bacteria. Nevertheless, further evaluation is needed to determine why *usp* is being co-transcribed with IS*Ftu1* only during early and mid log growth phase, and then both genes are transcribed individually during late log and stationary phase.

Nitric oxide production by SNAP resulted in the highest co-expression of the gene encoding the universal stress protein and IS*Ftu1*. It is possible that the co-expression of *usp* and IS*Ftu1* is important to maintain the viability of the bacteria by reducing the effects of SNAP via the

universal stress protein. Spermine also increased expression of the co-transcribed *usp* and IS*Ftu1*. Spermine is a polyamine that stabilizes secondary structures. Therefore, the increased abundance of co-transcribed *usp* and IS*Ftu1* could be due to increased stability of this transcript.

Since Usp is known to be involved with the bacteria's response to environmental stress (Siegele, 2005), it is interesting that this gene is located upstream of the IS*Ftu1* transposase and is co-transcribed with it. The universal stress protein is most likely promoting *F. tularensis* survival, while IS*Ftu1* may be contributing to the sequestration and/or stability of the co-transcribed mRNA. Carlson et. al. determined that a 30 bp region upstream of the majority of IS*Ftu1* genes in *F. tularensis* was conserved, so they cloned this region along with IS*Ftu1* and a reporter gene and then compared the expression levels of the reporter in response to spermine versus untreated LVS (Carlson et al., 2009). Since the expression of IS*Ftu1* expression increased in the presence of spermine compared to untreated LVS, they proposed that there were spermine-responsive elements in the IS*Ftu1* promoters inducing the expression of the downstream gene(s). Our study identified another possible role for the expression of an IS element, since the synthesis of IS*Ftu1* was downstream of a highly expressed gene and may contribute to transcript stability and/or the regulation of expression for the upstream gene.

The secondary RNA structures predicted that *usp* co-transcribed with IS*Ftu1* was more readily formed than when *usp* was transcribed alone. Collectively, our results suggest that the co-expression of *usp* with IS*Ftu1* may provide a mechanism to protect and sequester *usp*, preventing translation of this gene when conditions are not stressful and the need for both *usp* and IS*Ftu1* is minimal (i.e. during early and mid log growth phase). However, when nutrients become limiting (i.e. during late growth and stationary phase), monocistronic *usp* transcripts that have less structure are produced, providing easy access to the translational machinery. Additional study will provide a better understanding of *usp* function(s) and the contribution of IS elements to the differential expression and virulence observed for the various *F. tularensis* subpopulations.

REFERENCES

Akiyama, M., Horiuchi, T., Sekiguchi, M., 1987. Molecular cloning and nucleotide sequence of the mutT mutator of *Escherichia coli* that causes A:T to C:G transversion. Molecular & general genetics 206 (1), 9-16.

Andersson, H., Hartmanová, B., Kuolee, R., Rydén, P., Conlan, W., Chen, W., Sjöstedt, A., 2006. Transcriptional profiling of host responses in mouse lungs following aerosol infection with type A *Francisella tularensis* , Journal of medical microbiology 55 (3), 263-271.

Anthony, L.S., Morrissey, P.J., Nano, F.E., 1992. Growth inhibition of *Francisella tularensis* live vaccine strain by INF-gamma-activated macrophages is mediated by reactive nitrogen intermediates derived from L-arginine metabolism. Journal of immunology 148 (6), 1829-1834.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22, 195-201.

Asare, R., Abu Kwaik, Y., 2010. Exploitation of host cell biology and evasion of immunity by *Francisella tularensis*. Frontiers in Microbiology 1, 145. doi: 10.3389/fmicb.2010.00145.

Bandara, A.B., Champion, A.E., Wang, X., Berg, G., Apicella, M.A., McLendon, M.K., Azadi, P., Snyder, D.S., Inzana, T.J., 2011. Isolation and mutagenesis of a capsule-like complex (CLC) from *Francisella tularensis*, and contribution of the CLC to *F. tularensis* virulence in mice. PloS one 6 (4), e19003. doi: 10.1371/journal.pone.0019003.

Bardaji, L., Anorga, M., Jackson, R.W., Martinez-Bilbao, A., Yanguas-Casas, N., Murillo, J., 2011. Miniature transposable sequences are frequently mobilized in the bacterial plant pathogen *Pseudomonas syringae pv. phaseolicola*, PLoS One 6 (10), e25773. doi: 10.1371/journal.pone.0025773 [doi].

Barker, J.R., Chong, A., Wehrly, T.D., Yu, J.J., Rodriguez, S.A., Liu, J., Celli, J., Arulanandam, B.P., Klose, K.E., 2009. The *Francisella tularensis* pathogenicity island encodes a secretion system that is required for phagosome escape and virulence. Molecular microbiology 74 (6), 1459-1470.

Beauregard, A., Curcio, M.J., Belfort, M., 2008. The take and give between retrotransposable elements and their hosts, Annu Rev Genet 42, 587-617. doi: 10.1146/annurev.genet.42.110807.091549.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2013. GenBank, Nucleic Acids Research 41, D36-D42. doi: 10.1093/nar/gks1195.

Bergman, N.H., Passalacqua, K.D., Gaspard, R., Shetron-Rama, L.M., Quackenbush, J., Hanna, P.C., 2005. Murine macrophage transcriptional responses to *Bacillus anthracis*

infection and intoxication, Infection and immunity 73 (2), 1069-1080. doi: 10.1128/IAI.73.2.1069-1080.2005 10.

Bobay, L.M., Ochman, H., 2017. The evolution of bacterial genome architecture, Front Genet 8, 72. doi: 10.3389/fgene.2017.00072.

Bönemann, G., Pietrosiuk, A., Mogk, A., 2010. Tubules and donuts: a type VI secretion story. Molecular microbiology 76 (4), 815-821. doi: 10.1111/j.1365-2958.2010.07171.x.

Boyer, F., Fichant, G., Berthod, J., Vandenbrouck, Y., Attree, I., 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genomics 10, 104. doi: 10.1186/1471-2164-10-104.

Broms, J.E., Sjostedt, A., Lavander, M., 2010. The role of the *Francisella tularensis* pathogenicity island in type VI secretion, intracellular survival, and modulation of host cell signaling, Front Microbiol 1, 136. doi: 10.3389/fmicb.2010.00136.

Buchan, B., McCaffrey, R., Lindemann, S., Allen, L., Jones, B., 2009. Identification of migR, a regulatory element of the *Francisella tularensis* live vaccine strain *iglABCD* virulence operon required for normal replication and trafficking in macrophages, Infection and immunity 77 (6), 2517-2529. doi: 10.1128/IAI.00229-09.

Buchner, J., 2010. Bacterial Hsp90 - desperately seeking clients, Molecular microbiology 76 (3), 540-544. doi: 10.1111/j.1365-2958.2010.07140.x.

Busse, H.J., Huber, B., Anda, P., Escudero, R., Scholz, H.C., Seibold, E., Splettstoesser, W.D., Kampfer, P., 2010. Objections to the transfer of *Francisella novicida* to the subspecies rank of *Francisella tularensis* - response to Johansson et al, Int J Syst Evol Microbiol 60 (Pt 8), 1718-1720. doi: 10.1099/00207713-60-8-1718.

Carlson, P.J., Horzempa, J., O'Dee, D., Robinson, C., Neophytou, P., Labrinidis, A., Nau, G., 2009. Global transcriptional response to spermine, a component of the intramacrophage environment, reveals regulation of *Francisella* gene expression through insertion sequence elements. Journal of Bacteriology 191 (22), 6855-6865. doi: 10.1128/JB.00995-09.

Carvalho, C.L., Lopes de Carvalho, I., Ze-Ze, L., Nuncio, M.S., Duarte, E.L., 2014. Tularaemia: a challenging zoonosis, Comp Immunol Microbiol Infect Dis 37 (2), 85-96. doi: 10.1016/j.cimid.2014.01.002.

Casadevall, A., 2008. Evolution of intracellular pathogens, Annu Rev Microbiol 62, 19-33. doi: 10.1146/annurev.micro.61.080706.093305.

Centers for Disease Control and Prevention (CDC), Department of Health and Human Services (HHS), 2012. Possession, use, and transfer of select agents and toxins; biennial review. Final rule. Federal Register 77 (194), 61083-61115.

Chen, J., Petrov, A., Johansson, M., Tsai, A., O'Leary, S.E., Puglisi, J.D., 2014. Dynamic pathways of -1 translational frameshifting, Nature 512 (7514), 328-332. doi: 10.1038/nature13428.

Cherwonogrodzky, J.W., Knodel, M.H., Spence, M.R., 1994. Increased encapsulation and virulence of *Francisella tularensis* live vaccine strain (LVS) by subculturing on synthetic medium. Vaccine 12 (9), 773-775.

Chong, A., Wehrly, T.D., Nair, V., Fischer, E.R., Barker, J.R., Klose, K.E., Celli, J., 2008. The early phagosomal stage of *Francisella tularensis* determines optimal phagosomal escape and *Francisella* pathogenicity island protein expression, Infect Immun 76 (12), 5488-5499. doi: 10.1128/IAI.00682-08.

Clemens, D.L., Lee, B.Y., Horwitz, M.A., 2005. *Francisella tularensis* enters macrophages via a novel process involving pseudopod loops, Infect Immun 73 (9), 5892-5902.

Clemens, D.L., Lee, B.Y., Horwitz, M.A., 2004. Virulent and avirulent strains of *Francisella tularensis* prevent acidification and maturation of their phagosomes and escape into the cytoplasm in human macrophages. Infection and immunity 72, 3204-3217.

Croucher, N.J., Thomson, N.R., 2010. Studying bacterial transcriptomes using RNA-seq, Current opinion in microbiology 13 (5), 619-624. doi: 10.1016/j.mib.2010.09.009.

Darmon, E., Leach, D.R., 2014. Bacterial genome instability, Microbiol Mol Biol Rev 78 (1), 1-39. doi: 10.1128/MMBR.00035-13.

Darrah, P.A., Hondalus, M.K., Chen, O., Ischiropoulos, H., Mosser, D.M., 2000. Cooperation between reactive oxygen and nitrogen intermediates in killing of *Rhodococcus equi* by activated macrophages, Infection and Immunity 68 (6), 3587-3593.

Das, S., Chaudhuri, K., 2003. Identification of a unique IAHP (IcmF associated homologous proteins) cluster in *Vibrio cholerae* and other proteobacteria through in silico analysis. In Silico Biology 3 (3), 287-300.

Dennis, D.T., Inglesby, T.V., Henderson, D.A., Bartlett, J.G., Ascher, M.S., Eitzen, E., Fine, A.D., Friedlander, A.M., Hauer, J., Layton, M., Lillibridge, S.R., McDade, J.E., Osterholm, M.T., O'Toole, T., Parker, G., Perl, T.M., Russell, P.K., Tonat, K., Working Group on Civilian Biodefense, 2001. Tularemia as a biological weapon: medical and public health management, JAMA 285 (21), 2763-2773.

Derbyshire, K.M., Kramer, M., Grindley, N.D., 1990. Role of instability in the cis action of the insertion sequence IS903 transposase, Proc Natl Acad Sci U S A 87 (11), 4048-4052.

Dobrindt, U., Hacker, J., 2001. Whole genome plasticity in pathogenic bacteria, Curr Opin Microbiol 4 (5), 550-557.

Ellis, J., Oyston, P.C., Green, M., Titball, R.W., 2002. Tularemia, Clin Microbiol Rev 15 (4), 631-646.

Escoubas, J.M., Prere, M.F., Fayet, O., Salvignol, I., Galas, D., Zerbib, D., Chandler, M., 1991. Translational control of transposition activity of the bacterial insertion sequence IS1, EMBO J 10 (3), 705-712.

Farabaugh, P.J., 1996. Programmed translational frameshifting, Microbiol Rev 60 (1), 103-134.

Farlow, J., Wagner, D.M., Dukerich, M., Stanley, M., Chu, M., Kubota, K., Petersen, J., Keim, P., 2005. *Francisella tularensis* in the United States, Emerg Infect Dis 11 (12), 1835-1841. doi: 10.3201/eid1112.050728.

Feldman, K.A., Enscore, R.E., Lathrop, S.L., Matyas, B.T., McGuill, M., Schriefer, M.E., Stiles-Enos, D., Dennis, D.T., Petersen, L.R., Hayes, E.B., 2001. An outbreak of primary pneumonic tularemia on Martha's Vineyard, N Engl J Med 345 (22), 1601-1606. doi: 10.1056/NEJMoa011374.

Filloux, A., 2009. The type VI secretion system: a tubular story. The EMBO journal 28 (4), 309-310. doi: 10.1038/emboj.2008.301.

Foley, J.E., Nieto, N.C., 2010. Tularemia, Vet Microbiol 140 (3-4), 332-338. doi: 10.1016/j.vetmic.2009.07.017.

Forslund, A.L., Kuoppa, K., Svensson, K., Salomonsson, E., Johansson, A., Byström, M., Oyston, P.C., Michell, S.L., Titball, R.W., Noppa, L., Frithz-Lindsten, E., Forsman, M., Forsberg, A., 2006. Direct repeat-mediated deletion of a type IV pilin gene results in major virulence attenuation of *Francisella tularensis*. Molecular microbiology 59 (6), 1818-1830.

Forslund, A.L., Salomonsson, E.N., Golovliov, I., Kuoppa, K., Michell, S., Titball, R., Oyston, P., Noppa, L., Sjöstedt, A., Forsberg, A., 2010. The type IV pilin, PilA, is required for full virulence of *Francisella tularensis* subspecies *tularensis*. BMC Microbiology 10, 227. doi: 10.1186/1471-2180-10-227.

Fowler, R.G., Schaaper, R.M., 1997. The role of the *mutT* gene of *Escherchia coli* in maintaining replication fidelity, FEMS Microbiology Reviews 21 (1), 43-54. doi: 10.1111/j.1574-6976.1997.tb00344.x.

Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution, Nat Rev Microbiol 3 (9), 722-732.

Geissler, E., 2005. Alibek, tularemia and the battle of Stalingrad, Chemical and Biological Weapons Convention Bulletin 69, 10-15.

Giannoukos, G., Ciulla, D.M., Huang, K., Haas, B.J., Izard, J., Levin, J.Z., Livny, J., Earl, A.M., Gevers, D., Ward, D.V., Nusbaum, C., Birren, B.W., Gnirke, A., 2012. Efficient and

robust RNA-seq process for cultured bacteria and complex community transcriptomes, Genome Biology 13 (3), R23. doi: 10.1186/gb-2012-13-3-r23.

Gil, H., Benach, J., Thanassi, D., 2004. Presence of pili on the surface of *Francisella tularensis*, Infection and immunity 72, 3042-3047.

Golovliov, I., Sjöstedt, A., Mokrievich, A., Pavlov, V., 2003. A method for allelic replacement in *Francisella tularensis*. FEMS microbiology letters 222 (2), 273-280.

Grindley, N.D., Joyce, C.M., 1980. Genetic and DNA sequence analysis of the kanamycin resistance transposon Tn903, Proc Natl Acad Sci U S A 77 (12), 7176-7180.

Grudniak, A.M., Markowska, K., Wolska, K.I., 2015. Interactions of *Escherichia coli* molecular chaperon HtpG with DnaA replication initiator DNA, Cell Stress and Chaperones 20, 951-957. doi: 10.1007/s12192-015-0623-y.

Gunn, J.S., Ernst, R.K., 2007. The structure and function of *Francisella* lipopolysaccharide, Annals of the New York Academy of Sciences 1105, 202-218.

Hacker, J., Kaper, J.B., 2000. Pathogenicity islands and the evolution of microbes, Annu Rev Microbiol 54, 641-679. doi: 10.1146/annurev.micro.54.1.641.

Hadjiargyrou, M., Delihas, N., 2013. The intertwining of transposable elements and non-coding RNAs, International Journal of Molecular Sciences 14 (7), 13307-13328. doi: 10.3390/ijms140713307.

Hinnebusch, J., Tilly, K., 1993. Linear plasmids and chromosomes in bacteria, Mol Microbiol 10 (5), 917-922.

Huber, B., Escudero, R., Busse, H.J., Seibold, E., Scholz, H.C., Anda, P., Kämpfer, P., Splettstoesser, W.D., 2010. Description of *Francisella hispaniensis* sp. nov., isolated from human blood, reclassification of *Francisella novicida* (Larson et al. 1955) Olsufiev et al. 1959 as *Francisella tularensis* subsp. *novicida* comb. nov. and emended description of the genus *Francisella*. International Journal of Systematic and Evolutionary Microbiology 60 (8), 1887-1896. doi: 10.1099/ijs.0.015941-0.

Igarashi, K., Kashiwagi, K., 2000. Polyamines: mysterious modulators of cellular functions, Biochemical and Biophysical Research Communications 271 (3), 559-564. doi: 10.1006/bbrc.2000.2601.

Johansson, A., Celli, J., Conlan, W., Elkins, K.L., Forsman, M., Keim, P.S., Larsson, P., Manoil, C., Nano, F.E., Petersen, J.M., Sjostedt, A., 2010. Objections to the transfer of *Francisella novicida* to the subspecies rank of *Francisella tularensis*, Int J Syst Evol Microbiol 60 (Pt 8), 1717-8; author reply 1718-20. doi: 10.1099/ijs.0.022830-0.

Kammler, M., Schön, C., Hantke, K., 1993. Characterization of the ferrous iron uptake system of *Escherichia coli.*, Journal of Bacteriology 175 (19), 6212-6219.

Kawula, T.H., Hall, J.D., Fuller, J.R., Craven, R.R., 2004. Use of transposon-transposase complexes to create stable insertion mutant strains of *Francisella tularensis* LVS. Applied and Environmental Microbiology 70 (11), 6901-6904.

Keim, P.S., Wagner, D.M., 2009. Humans and evolutionary and ecological forces shaped the phylogeography of recently emerged diseases, Nat Rev Microbiol 7 (11), 813-821. doi: 10.1038/nrmicro2219.

Kingry, L.C., Petersen, J.M., 2014. Comparative review of *Francisella tularensis* and *Francisella novicida*, Frontiers in Cellular and Infection Microbiology 4, 35. doi: 10.3389/fcimb.2014.00035.

Kodaira, M., Biswas, S.B., Kornberg, A., 1983. The *dnaX* gene encodes the DNA polymerase III holoenzyme tau subunit, percursor of the gamma subunit, the dnaZ gene product. Molecular & general genetics 192 (1-2), 80-86.

Kolstø, A., Tourasse, N., Okstad, O., 2009. What sets *Bacillus anthracis* apart from other *Bacillus* species? Annual Review of Microbiology 63 (1), 451-476. doi: 10.1146/annurev.micro.091208.073255.

Kovacs-Simon, A., Titball, R.W., Mitchell, S.L., 2011. Lipoproteins of bacterial pathogens, Infection and immunity 79 (2), 548-561. doi: 10.1128/IAI.00682-10.

Kuan, C.T., Liu, S.K., Tessman, I., 1991. Excision and transposition of Tn5 as an SOS activity in *Escherichia coli*, Genetics 128 (1), 45-57.

Kusumoto, M., Ooka, T., Nishiya, Y., Ogura, Y., Saito, T., Sekine, Y., Iwata, T., Akiba, M., Hayashi, T., 2011. Insertion sequence-excision enhancer removes transposable elements from bacterial genomes and induces various genomic deletions, Nat Commun 2, 152.

Lambowitz, A., Zimmerly, S., 2011. Group II introns: mobile ribozymes that invade DNA, Cold Spring Harbor perspectives in biology 3 (8), a003616. doi: 10.1101/cshperspect.a003616.

Larson, M.A., Fey, P.D., Bartling, A.M., Iwen, P.C., Dempsey, M.P., Francesconi, S.C., Hinrichs, S.H., 2011. *Francisella tularensis* molecular typing using differential insertion sequence amplification, J Clin Microbiol 49 (8), 2786-2797. doi: 10.1128/JCM.00033-11.

Larson, M.A., Nalbantogul, U., Sayood, K., Zentz, E.B., Bartling,A.M., Francesconi,S.C., Fey, P.D., Dempsey, M.P., Hinrichs, S.H., 2015. *Francisella tularensis* subtype A.II genomic plasticity in comparison with subtype A.I, PloS one 10 (4), e0124906. doi: 10.1371/journal.pone.0124906.

Larsson, P., Elfsmark, D., Svensson, K., Wikstrom, P., Forsman, M., Brettin, T., Keim, P., Johansson, A., 2009. Molecular evolutionary consequences of niche restriction in *Francisella tularensis*, a facultative intracellular pathogen, PLoS Pathog 5 (6), e1000472. doi: 10.1371/journal.ppat.1000472.

Leiman, P.G., Basler, M., Ramagopal, U.A., Bonanno, J.B., Sauder, J.M., Pukatzki, S., Burley, S.K., Almo, S.C., Mekalanos, J.J., 2009. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proceedings of the National Academy of Sciences of the United States of America 106 (11), 4154-4159. doi: 10.1073/pnas.0813360106.

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., Lopez, R., 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Research 43, W580-W584.

Lindemann, S.R., Peng, K., Long, M.E., Hunt, J.R., Apicella, M.A., Monack, D.M., Allen, L.A., Jones, B.D., 2011. *Francisella tularensis* Schu S4 O-antigen and capsule biosynthesis gene mutants induce early cell death in human macrophages. Infection and immunity 79 (2), 581-594. doi: 10.1128/IAI.00863-10.

Lohe, A.R., De Aguiar, D., Hartl, D.L., 1997. Mutations in the mariner transposase: the D,D(35)E consensus sequence is nonfunctional, Proc Natl Acad Sci U S A 94 (4), 1293-1297.

Ma, L.S., Lin, J.S., Lai, E.M., 2009. An IcmF family protein, ImpLM, is an integral inner membrane protein interacting with ImpKL, and its walker a motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*. Journal of Bacteriology 191 (13), 4316-4329. doi: 10.1128/JB.00029-09.

MacMicking, J., Xie, Q., Nathan, C., 1997. Nitric oxide and macrophage function, Annual Review of Immunology 15, 323-350. doi: 10.1146/annurev.immunol.15.1.323.

Magnusson, M., Tobes, R., Sancho, J., Pareja, E., 2007. Cutting edge: natural DNA repetitive extragenic sequences from gram-negative pathogens strongly stimulate TLR9, J Immunol 179 (1), 31-35.

Mahillon, J., Chandler, M., 1998. Insertion sequences, Microbiol Mol Biol Rev 62 (3), 725-774.

Maier, T.M., Casey, M.S., Becker, R.H., Dorsey, C.W., Glass, E.M., Maltsev, N., Zahrt, T.C., Frank, D.W., 2007. Identification of *Francisella tularensis* Himar1-based transposon mutants defective for replication in macrophages. Infection and immunity 75 (11), 5376-5389.

Maki, H., Sekiguchi, M., 1992. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. Nature 355 (6357), 273-275.

Mc Gann, P., Rozak, D.A., Nikolich, M.P., Bowden, R.A., Lindler, L.E., Wolcott, M.J., Lathigra, R., 2010. A novel brain heart infusion broth supports the study of common *Francisella tularensis* serotypes, J Microbiol Methods 80 (2), 164-171. doi: 10.1016/j.mimet.2009.12.005 [doi].

McLendon, M.K., Apicella, M.A., Allen, L.A., 2006. *Francisella tularensis*: taxonomy, genetics, and immunopathogenesis of a potential agent of biowarfare, Annu Rev Microbiol 60, 167-185. doi: 10.1146/annurev.micro.60.080805.142126.

Meibom, K.L., Forslund, A.L., Kuoppa, K., Alkhuder, K., Dubail, I., Dupuis, M., Forsberg, A., Charbit, A., 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in *Francisella tularensis*, Infect Immun 77 (5), 1866-1880. doi: 10.1128/IAI.01496-08.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordoñez, C.L., Lory, S., Walz, T., Joachimiak, A., Mekalanos, J.J., 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. Science 312 (5579), 1526-1530.

Muñoz-López, M., García-Pérez, J., 2010. DNA transposons: nature and applications in genomics, Current genomics 11 (2), 115-128. doi: 10.2174/138920210790886871.

Nam, H.J., Jeon, J., Kim, S., 2009. Bioinformatic approaches for the structure and function of membrane proteins, BMB Reports 42 (11), 697-704.

Nano, F.E., Zhang, N., Cowley, S.C., Klose, K.E., Cheung, K.K., Roberts, M.J., Ludu, J.S., Letendre, G.W., Meierovics, A.I., Stephens, G., Elkins, K.L., 2004. A *Francisella tularensis* pathogenicity island required for intramacrophage growth, J Bacteriol 186 (19), 6430-6436. doi: 10.1128/JB.186.19.6430-6436.2004.

Newbury, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D., Higgins, C.F., 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences, Cell 48 (2), 297-310.

Newstead, S.L., Gates, A.J., Hartley, M.G., Rowland, C.A., Williamson, E.D., Lukaszewski, R.A., 2014. Control of intracellular *Francisella tularensis* by different cell types and the role of nitric oxide, Journal of Immunology Research 2014, 694717. doi: 10.1155/2014/694717.

Nierman, W.C., DeShazer, D., Kim, H.S., Tettelin, H., Nelson, K.E., Feldblyum, T., Ulrich, R.L., Ronning, C.M., Brinkac, L.M., Daugherty, S.C., Davidsen, T.D., Deboy, R.T., Dimitrov, G., Dodson, R.J., Durkin, A.S., Gwinn, M.L., Haft, D.H., Khouri, H., Kolonay, J.F., Madupu, R., Mohammoud, Y., Nelson, W.C., Radune, D., Romero, C.M., Sarria, S., Selengut, J., Shamblin, C., Sullivan, S.A., White, O., Yu, Y., Zafar, N., Zhou, L., Fraser, C.M., 2004. Structural flexibility in the *Burkholderia mallei* genome, Proc Natl Acad Sci U S A 101 (39), 14246-14251. doi: 10.1073/pnas.0403306101.

Nystrom, T., Neidhardt, F.C., 1992. Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in *Escherichia coli*, Molecular microbiology 6 (21), 3187-3198.

Overington, J.P., Al-Lazikani, B., Hopkins, A.L., 2006. How many drug targets are there? Nature Reviews Drug Discovery 5, 993-996. doi: 10.1038/nrd2199.

Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T., Prentice, M.B., Sebaihia, M., James, K.D., Churcher, C., Mungall, K.L., Baker, S., Basham, D., Bentley, S.D., Brooks, K., Cerdeno-Tarraga, A.M., Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Karlyshev, A.V., Leather, S., Moule, S., Oyston, P.C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrell, B.G., 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague, Nature 413 (6855), 523-527. doi: 10.1038/35097083.

Perkins-Balding, D., Duval-Valentin, G., Glasgow, A.C., 1999. Excision of IS492 requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*, J Bacteriol 181 (16), 4937-4948.

Plasterk, R.H., Izsvak, Z., Ivics, Z., 1999. Resident aliens: the Tc1/mariner superfamily of transposable elements, Trends Genet 15 (8), 326-332.

Pollitzer, R., 1967. History and Incidence of Tularemia in the Soviet Union, a Review. Institute for Contemporary Russian Studies, Fordham University, Bronx, N.Y.

Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., Mekalanos, J.J., 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proceedings of the National Academy of Sciences of the United States of America 104 (39), 15508-15513.

Pukatzki, S., McAuley, S.B., Miyata, S.T., 2009. The type VI secretion system: translocation of effectors and effector-domains. Current opinion in microbiology 12 (1), 11-17. doi: 10.1016/j.mib.2008.11.010.

Qin, A., Mann, B.J., 2006. Identification of transposon insertion mutants of *Francisella tularensis tularensis* strain Schu S4 deficient in intracellular replication in the hepatic cell line HepG2. BMC Microbiology 6, 69.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer, Nature Biotechnology 29 (1), 24-26. doi: 10.1038/nbt.1754.

Rohmer, L., Brittnacher, M., Svensson, K., Buckley, D., Haugen, E., Zhou, Y., Chang, J., Levy, R., Hayden, H., Forsman, M., Olson, M., Johansson, A., Kaul, R., Miller, S.I., 2006. Potential source of *Francisella tularensis* Live Vaccine Strain attenuation determined by genome comparison, Infection and immunity 74 (12), 6895-6906. doi: 10.1128/IAI.01006-06.

Rowe, H.M., Huntley, J.F., 2015. From the outside-in: the *Francisella tularensis* envelope and virulence. Frontiers in Cellular and Infection Microbiology 5, 94. doi: 10.3389/fcimb.2015.00094.

Russo, B.C., Horzempa, J., O'Dee, D.M., Schmitt, D.M., Brown, M.J., Carlson Jr., P.E., Xavier, R.J., Nau, G.J., 2011. A *Francisella tularensis* locus required for spermine responsiveness is necessary for virulence, Infection and immunity 79 (9), 3665-3676. doi: 10.1128/IAI.00135-11.

Safi, H., Barnes, P.F., Lakey, D.L., Shams, H., Samten, B., Vankayalapati, R., Howard, S.T., 2004. IS6110 functions as a mobile, monocyte-activated promoter in *Mycobacterium tuberculosis*, Mol Microbiol 52 (4), 999-1012. doi: 10.1111/j.1365-2958.2004.04037.x.

Salomonsson, E.N., Forslund, A.L., Forsberg, A., 2011. Type IV pili in *Francisella* - a virulence trait in an intracellular pathogen, Frontiers in microbiology 2, 1-7. doi: 10.3389/fmicb.2011.00029.

Sandström, G., Löfgren, S., Tärnvik, A., 1988. A capsule-deficient mutant of *Francisella tularensis* LVS exhibits enhanced sensitivity to killing by serum but diminished sensitivity to killing by polymorphonuclear leukocytes, Infection and immunity 56 (5), 1194-1202.

Santic, M., Al-Khodor, S., Abu Kwaik, Y., 2010. Cell biology and molecular ecology of *Francisella tularensis*, Cell Microbiol 12 (2), 129-139. doi: 10.1111/j.1462-5822.2009.01400.x.

Schlieker, C., Zentgraf, H., Dersch, P., Mogk, A., 2005. ClpV, a unique Hsp100/Clp member of pathogenic proteobacteria, Biological Chemistry 386 (11), 1115-1127.

Seiler, N., 1999. 5 POLYAMINES AND THE IMMUNE SYSTEM, Polyamines in Health and Nutrition , 65.

Shao, H., Tu, Z., 2001. Expanding the diversity of the IS630-Tc1-mariner superfamily: discovery of a unique DD37E transposon and reclassification of the DD37D and DD39D transposons, Genetics 159 (3), 1103-1115.

Siegele, D.A., 2005. Universal stress proteins in *Escherichia coli*, Journal of Bacteriology 187 (18), 6253-6254. doi: 10.1128/JB.187.18.6253-6254.2005.

Sievers, F., Higgins, D., 2014. Clustal omega. Current protocols in bioinformatics 48, 1.25.1-1.25.33. doi: 10.1002/0471250953.bi0313s48.

Sjöstedt, A., 2007. Tularemia: history, epidemiology, pathogen physiology, and clinical manifestations, Ann N Y Acad Sci 1105 (1), 1-29. doi: 10.1196/annals.1409.009.

Skerman, V.B.D., McGowen, V., Sneath, P.H.A., 1980. Approved lists of bacterial names, International Journal of Systematic and Evolutionary Microbiology 30, 225-420.

Streit, S., Christoph, M.W., Erkan, M., Kleeff, J., Helmut, F., 200p. Northern blot analysis for detection and quatification of RNA in pancreatic cancer cells and tissues, Nature protocols 4 (1), 37-43. doi: 10.1038/nprot.2008.216.

Su, J., Yang, J., Zhao, D., Kawula, T.H., Banas, J.A., Zhang, J.R., 2007. Genome-wide identification of *Francisella tularensis* virulence determinants. Infection and Immunity 75 (6), 3089-3101.

Tajiri, T., Maki, H., Sekiguchi, M., 1995. Functional cooperation of MutT, MutM, and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli*. Mutation research 336 (3), 257-267.

Tancred, L., Telepnev, M.V., Golovliov, I., Andersson, B., Andersson, H., Lindgren, H., Sjöstedt, A., 2011. Administration of a nitric oxide donor inhibits mglA expression by intracellular *Francisella tularensis* and counteracts phagosomal escape and subversion of TNF-a, Journal of medical microbiology 60 (Pt 11), 1570-1583. doi: 10.1099/jmm.0.032870-0.

Thomason, M.K., Storz, G., 2010. Bacterial antisense RNAs: how many are there, and what are they doing? Annu Rev Genet 44, 167-188. doi: 10.1146/annurev-genet-102209-163523.

Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-perfomrance genomics data visualization and exploration, Briefings in Bioinformatics 14 (2), 178-192. doi: 10.1093/bib/bbs017.

Tourasse, N., Kolstø, A., 2008. Survey of group I and group II introns in 29 sequenced genomes of the *Bacillus cereus* group: insights into their spread and evolution, Nucleic Acids Research 36 (14), 4529-4548. doi: 10.1093/nar/gkn372.

Trevisanato, S.I., 2007a. The biblical plague of the Philistines now has a name, tularemia, Med Hypotheses 69 (5), 1144-1146.

Trevisanato, S.I., 2007b. The 'Hittite plague', an epidemic of tularemia and the first record of biological warfare, Med Hypotheses 69 (6), 1371-1374.

Trevisanato, S.I., 2004. Did an epidemic of tularemia in Ancient Egypt affect the course of world history? Med Hypotheses 63 (5), 905-910.

Wagner, E., Altuvia, S., Romby, P., 2002. Antisense RNAs in bacteria and their genetic elements, Advances in Genetics 46 (1), 361-398.

Wang, H., Smith, M.C., Mullany, P., 2006. The conjugative transposon Tn5397 has a strong preference for integration into its *Clostridium difficile* target site, J Bacteriol 188 (13), 4871-4878.

Wehrly, T.D., Chong, A., Virtaneva, K., Sturdevant, D.E., Child, R., Edwards, J.A., Brouwer, D., Nair, V., Fischer, E.R., Wicke, L., Curda, A.J., Kupko, J.J.,3rd, Martens, C., Crane, D.D., Bosio, C.M., Porcella, S.F., Celli, J., 2009. Intracellular biology and virulence determinants of *Francisella tularensis* revealed by transcriptional profiling inside macrophages, Cell Microbiol 11 (7), 1128-1150. doi: 10.1111/j.1462-5822.2009.01316.x.

Weinreich, M.D., Makris, J.C., Reznikoff, W.S., 1991. Induction of the SOS response in *Escherichia coli* inhibits Tn5 and IS50 transposition, J Bacteriol 173 (21), 6910-6918.

Yoshida, M., Kashiwagi, K., Shigemasa, A., Taniguchi, S., Yamamoto, K., Makinoshima, A., Ishiama, A., Igarashi, K., 2004. A unifying model for the role of polyamines in bacterial cell growth, the polyamine modulon. Journal of Biological Chemisty 279 (46008), 46013.

Zheng, J., Leung, K.Y., 2007. Dissection of a type VI secretion system in *Edwardseilla tarda*, Molecular microbiology 66, 1192-1206.