CHARACTERIZATION OF *MAP3773C,* FERRIC UPTAKE REGULATOR PROTEIN, IN IRON METABOLISM OF *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

By

Fernanda Miyagaki Shoyama

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Comparative Medicine and Integrative Biology – Doctor of Philosophy

2020

ABSTRACT

CHARACTERIZATION OF *MAP3773C,* FERRIC UPTAKE REGULATOR PROTEIN, IN IRON METABOLISM OF *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS*

By

Fernanda Miyagaki Shoyama

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of Johne's disease (JD) in ruminants, a chronic inflammation of the intestines characterized by persistent diarrhea that leads to malnutrition and muscular wasting (Rathnaiah *et al.*, 2017). Unfortunately to date, reliable JD diagnostics are lacking. Culture of MAP from feces has been the most reliable method for diagnosis of JD, however MAP requires eight to sixteen weeks to produce colonies in culture, presenting a major hurdle to diagnosis (Bannantine *et al.,* 2002). Currently, no current vaccine can protect animals against JD or prevent shedding of MAP (Garcia and Shalloo, 2015).

These challenges establish a need for a better understanding of the *MAP* physiology during infection. Unlike other mycobacteria, MAP has a unique requirement for supplementation of an iron-binding siderophore (mycobactin J) for optimal growth in laboratory media. A whole genome sequence of *MAP* K10 revealed a truncation of the *mbtA* gene that was speculated to have led to its mycobactin dependency (Li *et al.,* 2005). Zhu et al (2008) showed that MAP is able to transcribe all mycobactin synthesis genes in an intra-macrophage environment. Furthermore, several genes responsible for iron acquisition in infected tissues, including genes responsible for mycobactin biosynthesis have been shown to be upregulated in naturally infected tissues (Janagama *et al.,* 2010). Its known that iron plays an important role in vital biological processes; however high intracellular concentration of free iron can lead to toxicity to the bacteria. As such, bacteria activate expression of specific group of genes that are controlled by a

metal-sensing regulatory transcription factor. In 2009, Janagama and others identified and characterized the iron dependent regulator (IdeR) in MAP. In addition, in a MAP-specific genomic island, MAP carries three putative ferric uptake regulator (Fur) boxes, an iron regulated transcriptional control motif, (Stratmann *et al.,* 2004). To date, nothing is known about the role of Fur in MAP. To elucidate the mechanisms of iron homeostasis in MAP*,* we investigated LSP15 using a transposon mutant *MAP3776c*. We demonstrated a phenotype for LSP15 genes in a culture model of cell entry and survival and suggest a function in epithelial cell pathogenesis, however further functional analysis with complementation by a MAP3776c::Tn strain would be necessary to confirm these findings. Additionally, full characterization of a Fur-like protein (*MAP3773c*) was performed. Using PRODORIC for computational analysis, 23 different pathways that were likely regulated by *MAP3773c* were identified. These findings were confirmed using a chromatin immunoprecipitation assay followed by high-throughput sequencing (ChIP-seq), that revealed 58 regions where Fur binds under iron replete and deplete conditions. From those, three were directed related to iron regulation *MAP3638c* (hemophore-like protein), *MAP3736c* (Fur box) and *MAP3776c* (ABC transporter*).* Using the Fur box consensus sequence, we confirmed binding specificity and Mn^{2+} availability by a chemiluminescent electrophoresis mobility shift assay (EMSA). A transcriptional profile of the parent MAP K10, deletion mutant of *MAP3773c* and the complemented strains was developed under iron replete and depleted conditions. However, under the current experimental conditions, we are unable to conclude if the lack of transcriptional responses in our study was indicative of a lack of FUR activity.

This dissertation is dedicated to the memory of my mom, Lourdes Hitomi Miyagaki.
Thank you for all your sacrifice, for all you did for me, and all you taught me so I could become the person I am today.
And to my dad Wilson Nobuaki Shoyama,
Thank you for always believing in me and for always encouraging me to follow my dreams

ACKNOWLEDGEMENTS

My sincerest gratitude to my mentor and friend Dr. Srinand Sreevatsan. First for believing in me as a Scientist, but the most importantly, for always challenging me to be a critical thinker. Thank you for your guidance throughout my time in your lab as a technician and as grad student, always encouraging me to keep going, especially when experiments failed. I am also thankful for all the opportunities and doors that you opened for me, always providing me ways to improve my skills as a scientist. Finally, thank you for sharing your expertise and knowledge.

I would like to thank Dr. John Bannantine for having me in his laboratory at USDA, and spending time at the bench with me. In addition, I am appreciative for his numerous data discussions on the phone.

I am deeply grateful to Dr. Raul Barletta for the construction of the Fur mutant, for discussions on the manuscript, and for always being available to help and support during this journey.

I thank the Comparative Medicine and Integrative Biology Grad Program at Michigan State. I appreciate Dr Colleen Hegg and Dimity Palazzola for all the support.

I thank the Veterinary Medicine Grad Program at the University of Minnesota. I appreciate Dr Mark Rutherford and Dr Michael Murtaugh for all the support

I thank my committee members: Dr. Linda Mansfield, Dr. Paul Coussens and Dr. Rob Abramovich for their advice on my research. I also appreciate Dr. Rob Abramovitch for teaching me new techniques and opening his lab to me to help my work progress.

I acknowledge all the personnel from Sreevatsan lab since Minnesota. They include Dr. Elise Lamont for teaching me mycobacterial techniques, Dr. Sylvia Wanzala for always being kind and making sure we were always calm, and Anum Hadi for being the warrior with me when we

moved to Michigan. I especially want to thank my friend Evan P. Brenner, not only for all the science discussion and experiment brainstorming, but also for always being my partner in crime in the lab and supporting me in this crazy grad student life.

I acknowledge all the funding support I received throughout my PhD, MnDrive Global Food Ventures, USDA, Veterinary Population Medicine – University of Minnesota and CMIB Program.

I thank my friends that Grad School gave me, Catalina, Irene, Whitney and Elaine, for being the fun, relaxing part of the PhD.

I am extremely grateful for the unwavering support, patience, and love of my boyfriend Nathanael McNabb, always showing me the best in life and how to be a better person, but most importantly, making sure that I could see the world outside the grad student life.

Finally, I would like to thank ALL my family. I am 100% sure this dissertation would not have been possible without their support; my parents and my grandparents for the love and endless support on education; my second moms: Mitie, Miuki, Yumie and Kaori, no words can express what you have done for me - thank you for always being there. Thank for my sisters, Leticia and Marina, for the virtual hugs and for always being the best cheerleaders.

Lastly, thank you for my best friend, Jack, that started this journey with me and made my life so much happier. In addition, thank Jack and Breezy for the pure love you have always given me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Functional characterization of iron regulation by *Mycobacterium avium* subsp. *paratuberculosis* (MAP) has revealed that MAP employs a complex protein pathway in response to iron starvation *in vitro* or *ex vivo*. Previous work demonstrated that the iron-dependent repressor protein (IdeR) not only controls mycobactin synthesis genes inside macrophages (Zhu *et al.,* 2008) but also regulates transcription of genes involved in iron acquisition (*mbtB*), iron storage (*bfrA*) and oxidative stress (Janagama *et al.,* 2009).

While expression of the *ideR* regulon varies with the concentration of intrabacterial iron, there seemed to be an external metal-sensing regulator missing for iron scavenging. It is known that MAP's genome contains 6 unique genomic islands that are absent in other mycobacteria (Alexander *et al.,* 2009). Transcriptional analysis by our laboratory suggests that two of these large sequence polymorphisms, LSP14 and LSP15, form a major genomic island that may play a role in iron homeostasis (Janagama *et al.,* 2010). However, it is not clear whether the genes encoded in those MAP-specific LSPs contribute to the pathogenicity of MAP (Meißner *et al.,* 2014). In 2014, Eckelt and others confirmed that LSP14 and LSP15 are involved in metal homeostasis. Furthermore, using a transposon mutant with a disruption in the LSP15 gene *MAP3776c*, Behr's group showed that when compared to the wild-type strain, the mutant demonstrated increased iron uptake as well as higher intracellular iron concentrations, indicating that LSP15 genes contribute to iron acquisition by MAP (Wang *et al.,* 2016). Corroborating the importance of LSP15 genes during infection, the same group showed that a disruption in *MAP3776c* reduces bacterial fitness by 10-fold in the mouse model (Wang *et al.,* 2014). By bioinformatic analysis, LSP15 is predicted to encode an ATP-binding cassette (ABC)

transporter, a metal uptake regulator and a gene that may be involved in cobalamin synthesis (Alexander *et al.,* 2009). Previous studies in our lab showed genes LSP15 were differentially regulated in infected bovine tissues, and interestingly, *MAP3773c*, a predicted ferric uptake regulator (fur) protein, was downregulated in the tissues and upregulated in experimentally infected macrophages (Janagama *et al.,* 2010). The same result was reported in the ABC transporter operon (LSP14) that contains two fur binding boxes (Stratmann *et al.,* 2004; Janagama *et al.,* 2010). Thus, it is already known that both LSP14 and LSP15 are involved in iron homeostasis, but there is no information on whether or not Fur contributes to this regulation. Originally described in *E. coli*, Fur is an iron-responsive repressor protein that controls the expression of genes for siderophore biosynthesis and iron transportation (Hantke, 1981). It has been reported that intracellular iron concentrations in many bacteria are under the control of Fur (Guerinot, 1994). During infection, in a metal-depleted environment, Fur allows for efficient acquisition of iron, leading to enhanced fitness of the pathogen (Lamont, Xu and Sreevatsan, 2013). Supporting this critical role of Fur in virulence, a deletion of *fur* in numerous bacterial pathogens most often resulted in partial or complete attenuation in animal models of infection. For activation of responsive genes, one Fe^{2+} ion per monomer enables Fur dimerization and binding to a specific 19bp DNA sequence (the "Fur box") that obstructs RNA polymerase and subsequent transcription (Lee and Helmann, 2007). Beside Fe(II), it has been reported in *B. subtilis* that binding of Mn(II) and Zn(II) can also activate Fur and are effective at triggering changes needed for DNA binding. Both metals are preferred and widely used in biochemical studies as the activating cofactor, since Fe(II) requires strict anaerobic conditions (Ma, Faulkner and Helmann, 2012). It is still unknown how MAP-Fur is activated.

The overall objective of the studies presented in this thesis was to elucidate the mechanism of iron homeostasis in *Mycobacterium avium* subsp. *paratuberculosis.*

We hypothesize that *MAP3773c* (Fur), a gene likely acquired by horizontal gene transfer, encodes a protein that orchestrates metabolic pathways required for optimal environmental sensing of iron, providing an alternative iron pathway.

Aim 1: Determine the Ferric Uptake Regulator (Fur) regulon in *Mycobacterium avium* subsp. *paratuberculosis.*

Aim 2: Identify putative functions of *MAP3773c* in iron homeostasis.

This thesis is organized into six chapters. *Chapter 2* gives a basic understanding of the importance of investigating MAP and a literature review on the involvement of iron and pathogenesis. *Chapter 3* provides characterization of LSP15 and *MAP3773c* phenotype using a transposon mutant *MAP3776c.* In this chapter, *in vitro* assays showed potential roles of *MAP3776c* in MAP cell invasion and further supports the reported involvement of LSP15 in intracellular survival. *Chapter 4* provides a characterization of the Fur regulon in MAP using ChIP-seq to identify the regulon under iron-replete and -depleted conditions *in vitro*. Sequences were analyzed using CLC Genomics Workbench 12. Confirmation of physical binding of *MAP3773c* to the Fur box was carried out by chemiluminescent EMSA, using a labeled Fur box and a recombinant MAP Fur-like protein. *Chapter 5* provides a comparative analysis of transcriptomes from MAP K10 (wild-type), fur deletion mutant and complemented strains under metal-replete and -depleted conditions. *Chapter 6* summarizes the findings presented in this thesis and proposes future studies.

REFERENCES

REFERENCES

Alexander DC, Turenne CY, Behr MA. Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. J Bacteriol. 2009; 191(3):1018-25. DOI: 10.1128/JB.01340-08

Eckelt, E., Jarek, M., Fromke, C., Meens, J., Goethe, R. (2014) Identification of a lineage specific zinc responsive genomic island in *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics. 15:1076. doi: 10.1186/1471-2164-15-1076

Guerinot ML. (1994). Microbial iron transport. Annual Rev. Microbiol. 48,743-772. DOI:10.1146/annurev.mi.48.100194.003523

Hantke, K. (1981). Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol Gen Genet. 182, 288-92. DOI:10.1007/bf00269672

Janagama, H.K., Lamont, E.A., George, S., Bannantine, J.P., Xu, W.W., Tu, Z.J., Wells, S.J., Schefers, J., Sreevatsan, S. (2010) Primary transcriptomes of *Mycobacterium avium* subsp. *paratuberculosis* reveal proprietary pathways in tissue and macrophages. BMC Genomics. 11, 561. doi: 10.1186/1471-2164-11-561.

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Lamont EA, Xu WW, Sreevatsan S. Host-*Mycobacterium avium* subsp. *paratuberculosis* interactome reveals a novel iron assimilation mechanism linked to nitric oxide stress during early infection. BMC Genomics. 2013; 14:694. DOI: 10.1186/1471-2164-14-694.

Lee, J.W., Helmann, J.D. Functional specialization within the Fur family of metalloregulators. Biometals. 2007; (3-4):485-99. DOI: 10.1007/s10534-006-9070-7

Li, L., Bannantine, J.P., Zhang, Q., Amonsin, A., May, B.J., Alt, D., Banerji, N., Kanjilal, S., Kapur, V. (2005) The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc. Natl. Acad. Sci. USA. 102,12344-12349. DOI:10.1073/pnas.0505662102

Ma, Z., Faulkner, M.J., Helmann, J.D. (2012). Origins of specificity and cross-talk in metal ion sensing by *Bacillus subtilis* Fur. Molecular Microbiology. 86, 1144-1155. DOI:10.1111/mmi.12049

Meißner T, Eckelt E, Basler T, Meens J, Heinzmann J, Suwandi A, Oelemann WMR, Trenkamp

S, Holst O, Weiss SF, Bunk B, Spröer C, Gerlach GF, Goethe R. (2014) The *Mycobacterium avium* ssp. *paratuberculosis* specific *mpt*D gene is required for maintenance of the metabolic homeostasis necessary for full virulence in mouse. Front. Cell. Infect. Microbiol. DOI: 10.3389/fcimb.2014.00110

Rathnaiah, G., Zinniel, D.K., Bannantine J.P., Stabel, J.R., Gröhn, Y.T., Collins, M.T., Barletta, R.G. (2017) Pathogenesis, Molecular Genetics, and Genomics of *Mycobacterium avium* subsp. *paratuberculosis*, the etiologic Agent of Johne's Disease. Front Vet Sci. 4:187. doi:10.3389/fvets.2017.00187

Stratmann, J., Strommenger, B., Goethe, R., Dohmann, K., Gerlach, G.F., Stevenson, K., Li, L., Zhang, Q., Kapur, V., Bull, T. (2004) A 38-kilobase pathogenicity island specific for *Mycobacterium avium* subsp. *paratuberculosis* encodes cell surface proteins expressed in the host. Infect. Immun. 72, 1265–1274. DOI: 10.1128/IAI.72.3.1265-1274.2004

Wang, J., Pritchard, J.R., Kreitmann, L., Montpetit, A., Behr, M.A. (2014) Disruption of *Mycobacterium avium* subsp. *paratuberculosis*-specific gene impairs in vivo fitness. BMC Genomics. 15:415. DOI:10.1186/1471-2164-15-415

Wang, J., Moolji, J., Dufort, A., Staffa, A., Domenech, P., Reed, M.B., Behr, M.A. (2016) Iron acquisition in *Mycobacterium avium* subsp. *paratuberculosis.* J. Bacteriol. 198:857–866. DOI:10.1128/JB.00922-15

Whittington, R., Donat, K., Weber, M.F., Kelton, D., Nielsen, S.S., Eisenber, S. (2019) Control of paratuberculosis: who, why and how. A review of 48 countries. BMC Veterinary Research. 15:198. DOI:10.1186/s12917-019-1943-4

Zhu, X., Tu, Z.J., Coussens, P.M., Kapur, V., Janagama, H., Naser, S., Sreevatsan, S. (2008) Transcriptional analysis of diverse strains *Mycobacterium avium* subspecies *paratuberculosis* in primary bovine monocyte derived macrophages. Microbes and Infection. 1274-1282. DOI:10.1016/j.micinf.2008.07.025

CHAPTER TWO: LITERATURE REVIEW

2.1. *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS* (MAP)

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of paratuberculosis, also known as Johne's disease (JD), a chronic enteritis affecting ruminants like cattle, sheep, and goats. Advanced stages of the disease are characterized by a malabsorption syndrome that leads to diarrhea, submandibular edema, malnutrition, anemia, lethargy, muscular wasting and death (Garcia and Shalloo, 2015). MAP is shed in feces of asymptomatic sub-clinical and/or clinically infected animals, establishing infection in new hosts via the fecal-oral route (Whittington and Sergeant, 2001). Primary infections occur *in utero* or during calfhood by direct contact with MAP via contaminated teats, colostrum/milk, pasture, feed, soil, and water (Windsor and Whittington, 2010). Disease detection is a difficult part of JD management, as most animals remain sub-clinical for several years after infection (Rathnaiah *et al.,* 2017). However, animals in sub-clinical stages intermittently shed MAP, ranging from fewer than 10^3 CFU/g of feces, to millions or billions of bacteria into the environment (Garcia and Shalloo, 2015). Visible signs of advanced disease include weight loss and decreased milk production. At necropsy, thick and inflexible intestinal sections are observed. Histopathology reveals granuloma formation in the small intestine and regional lymph nodes (Bannantine and Bermudez, 2013).

2.1.1 MAP INVASION AND ESTABLISHMENT

The host's intestinal epithelium serves as the primary barrier against MAP. Upon ingestion, MAP translocates across the intestinal mucosa via microfold cells (M cells) or enterocytes located in the Peyer's patches of the small intestines (Bannantine and Bermudez, 2013).

Although both routes serve as entry for MAP, studies have shown that M cells are the primary target for MAP infection (Secott and Wu, 2004; Ponnusamy *et al.,* 2013). This predilection can be explained by a combination of the presence of fibronectin attachment proteins (FAPs) – a family of fibronectin-binding proteins in MAP's cell wall – and the high density of the fibronectin receptor (β1 integrins) on the luminal surface of M cells (Secott, Lin and Wu, 2001; Secott and Wu, 2004; Everman *et al.,* 2018). Secott and others showed that when expression of FAPs was attenuated in MAP, the selective invasion into M cells was completely eliminated (Secott and Wu, 2004). Whether the route of entry has any physiological significance or impact on the outcome of the disease is still unknown (Bermudez *et al.*, 2010).

After crossing the epithelial barrier, MAP encounters dendritic cells and subepithelial macrophages (Lei *et al.,* 2007; Bermudez *et al.,* 2010). Coussens and others showed that MAP may use intestinal dendritic cell invasion as a strategic maneuver, since the primary function of intestinal dendritic cells is to sample and present commensal bacteria through tight junctions to the gut-associated lymphoid tissue (Coussens *et al.,* 2010; Coussens *et al.,* 2020). Furthermore, in the lamina propria, MAP also interacts with macrophages. Macrophages play a critical role in the host-pathogen dynamic, serving not only as a central effector for mediating the destruction of MAP, but also as a protected environment for the survival, proliferation and dissemination of MAP. Upon entry into the macrophage phagosome, MAP activates host immune response and phagosomes undergo stepwise development towards late endosomes (Arsenault *et al.,* 2014). The end result of phagosome–lysosome fusion is the destruction of MAP and presentation of antigens to T cells via major histocompatibility complex (MHC) (Coussens *et al.,* 2020). However, most phagosomes containing MAP fail to mature, resulting in the blockage of phagosome–lysosome fusion (Cheville *et al.*, 2001; Brumell and Scidmore, 2007). This blockade is essential to

restricting recognition of the pathogen by the host, allowing MAP unabated growth and persistence (Coussens *et al.,* 2020). *In vitro* experiments with infected macrophages showed a downregulation of MHC molecule expression along with greater expression of IL-10 and transforming growth factor beta (TGFβ) (Weiss and Souza, 2005). Corroborating these findings, an upregulation of inhibitory cytokines is also reported in other mycobacterial diseases, such as tuberculosis, where patients with advanced disease produced higher levels of interleukin (IL)-10 and TGFβ (Redford *et al.*, 2011). Studies in *M. tuberculosis* and MAP have shown that both cytokines play an important role in limiting the bactericidal capabilities of macrophages (Redford *et al.*, 2011; Souza *et al.*, 2013).

Obligate intracellular pathogens must avoid premature host cell death by apoptosis and necrosis to allow replication and infection of neighboring cells (Magombedze *et al.,* 2017). Therefore, the interaction between MAP and host cell dictates disease outcome. The ability of pathogenic bacteria to overcome varied stressors imposed by the host is a major virulence determinant (Danelishvili *et al.,* 2010; Kabara and Coussens, 2012). Thus, deep investigations into the mechanisms MAP uses to survive and overcome hostile conditions can improve our *in vitro* culture methods and consequently yield better diagnostics.

2.1.2 ECONOMIC IMPACT

Johne's disease (JD) causes significant direct and indirect economic impacts in the US and worldwide. A survey published in 2007 showed that 68% of the US dairy herds have at least one cow infected with MAP (NAHMS). The economic impact of JD for a farmer will be dependent on the number of animals affected, infected, infectious (shedding bacteria), presenting clinical signs or in the sub-clinical stage (Garcia and Shalloo, 2015). Culling of animals that are shedding

MAP is one of the most efficient JD control strategies. This can impose substantial economic losses, not only by culling of the animal, but also through the increased costs of replacements and losses of genetics in the herds (Smith *et al.,* 2010; Whittington *et al.,* 2019). Additionally, JD imposes compounding economic consequences, predisposing animals to clinical mastitis and increasing susceptibility for developing lameness, digestive and respiratory diseases, reducing milk production, and impairing fertility (Raizman *et al.,* 2007; Smith *et al., 2010*; Garcia and Shalloo, 2015; Rossi *et al.,* 2017). An estimate of annual economic losses from MAP infections in US dairy herds is between $21 to $79 per cow (Garcia and Shalloo, 2015). The economic production losses caused by *MAP* go far beyond an infected farm, as indirect losses may also occur at trade markets, either at a domestic or international level (Kennedy *et al.,* 2017).

2.1.3 CONTROL STRATEGIES FOR PARATUBERCULOSIS

Unfortunately, the impact of JD goes beyond simple economic losses. As the disease progresses, animal health and welfare deteriorate. The chronic enteritis caused by MAP is characterized by diarrhea, progressive weight loss and muscle wasting despite a good appetite and normal body temperature (Garcia and Shalloo, 2015). In addition, there is a potential for public health impacts from MAP: studies have investigated its association with Crohn's disease and sources for human exposure to the bacterium (Waddell *et al.,* 2016; Barkema *et al.,* 2017). As no treatment or cure is available, an effective control strategy is needed to reduce the spread and prevalence of disease (Barkema *et al.,* 2017).

2.1.3.1 VACCINE

Currently, there are no approved vaccines for use against JD in the United States, although options are available in other countries. Mycopar (Boehringer Ingelheim), a whole-cell vaccine comprised of an oil suspension of heat-killed MAP strain-18, was sold until 2019. Beside failing to offer complete protection against JD, Mycopar also causes severe granulomatous inflammation at the site of injection that may become abscessed, causing major discomfort to the animal (Uzonna *et al.,* 2003). A second vaccine, Silirum (Zoetis Animal Health), contains heat-killed MAP 316F strain and is commercialized in Australia. Studies demonstrate Silirum is able to reduce prevalence of clinical disease, but the immune responses induced by Silirum lead to cross-reactivity with tuberculin skin testing and vaccinated animals test seropositive for bovine tuberculosis (Stringer *et al.,* 2013; APVMA, 2014). A third vaccine also produced by Zoetis is Gudair. It also contains a heat-killed 316F strain and was developed for the control of ovine Johne's disease. In a five-year study, despite seeing declines in prevalence of JD after vaccination, sheep flocks continued to shed MAP and were at risk of spreading MAP infection (Windsor *et al.,* 2014).

A paradigm in mycobacterial vaccine development is that an attenuated strain (devoid of critical survival-associated genes) would be safe and provide protection in terms of reduced disease, shedding of MAP and protect animals from new infection. With this in mind, a consortium of MAP scientists performed a three-phase vaccine candidate evaluation that provided great insights for future vaccination development (Bannantine *et al.,* 2014a). Using the bovine monocyte-derived macrophage (MDM) model and an apoptosis study in macrophages, in Phase I, investigators evaluated different strains of MAP carrying deletions of critical housekeeping or putative virulence genes (Bannantine *et al.,* 2014b; Lamont *et al.,* 2014). A total of eight mutants

were identified as attenuated in the *ex vivo* macrophage model and were enrolled in Phase II, where tissue colonization of C57/BL6 mice was evaluated in a vaccination-challenge model (Bannantine *et al.,* 2014b). In the last phase, five candidates were evaluated for their ability to reduce fecal shedding, tissue colonization and pathology in a baby goat model. One candidate was able to lower incidence and severity of infection (Hines *et al.,* 2014) in the baby goat model. Concluding, the authors suggested that all investigations should be done in a natural host for MAP, as mouse models do not consistently reproduce MAP infection and infected mice do not show any clinical signs (Bannantine *et al.,* 2014a).

An optimal vaccine candidate is predicted to yield better outcomes if it is designed with more than one gene deletion (Bannantine *et al.,* 2014a). Supporting this concept, macrophage infection assays with *M. tuberculosis* using a double knockout mutant (*fbpA/sapM*) showed greater attenuation and was more immunogenic than either single knockout (*fbpA* and *sapM*) in macrophages (Saikolappan *et al.,* 2012). In addition, the same study showed that mice vaccinated with the double knockout strain elicited a stronger immune response when compared to the single knockout strains (Saikolappan *et al.,* 2012).

In summary, there is no ideal vaccine that can completely prevent infection, and no options currently exist for US farmers. Research on vaccine development for Johne's disease has always been a challenge due to the high costs of experiments using natural host models, lack of a good standardized animal model for JD, and incomplete understanding of the immune correlates of protection (Bannantine *et al.,* 2014a).

2.1.3.2 DIAGNOSTIC TESTS

Since existing vaccines are impractical as a control strategy, early diagnosis of MAP infection remains critical to curtail its spread. Unfortunately, diagnostic tests for JD are also far from perfect. In general, samples can be analyzed by microscopy, through *MAP* isolation in culture media, or via identification of MAP DNA by PCR. Though acid-fast staining is simple and fast, it cannot differentiate between MAP and other acid-fast bacilli, making a second confirmatory test necessary. To date, MAP isolation by culture is considered the "gold standard" test for JD; however, as explained in detail in section 4 of this chapter, MAP grows extremely slowly, taking up to 8 weeks to form colonies on solid media and 8-16 weeks to grow in liquid. Isolation of MAP in culture media is labor-intensive, time-consuming, and expensive. Additionally, fecal samples need to be decontaminated prior to inoculation of mycobacterial culture medium (Foddai *et al.,* 2009; Rathnaiah *et al.,* 2017). Decontamination methods for slowly growing mycobacteria aim to destroy irrelevant and usually rapidly growing microorganisms, so prolonged incubation with specific antimicrobial agents in combination with sedimentation, centrifugation or filtration of the fecal samples are common methods to isolate MAP (Whittington, 2009). Comparatively, nucleic acid detection by PCR makes direct testing of fecal samples much faster. Furthermore, improvements are constantly being made to available methods, such as for real-time, nested and TaqMan PCRs. A majority of these assays are based upon the amplification of the multicopy (n = ~17) MAP IS*900* genomic element. Another multicopy gene (IS*Mav2*) has also been used.

Indirect tests are based on a host immune response having been mounted against MAP. Indirect ELISA, for example, uses antigen to screen for MAP-specific antibodies in serum or milk. Older tests include agar gel immunodiffusion (AGID) and the complement fixation test (CFT). The

main disadvantage of immune-based tests is that they lack specificity, and results may vary depending on the species, age, and the experimental protocol used (Whittington *et al.,* 2017).

To develop more sensitive diagnostic tests, studies have focused on the discovery of novel antigens by MAP proteomic analyses (He and De Buck, 2010). In order to identify a large fraction of potential antigen candidates in a cost-effective manner, and by making use of close phylogenetic relationships, investigators probed a commercially available whole proteome analysis microarray from *M. tuberculosis* with serum from MAP-infected animals in different stages of the disease (Bannantine *et al.,* 2017). Of the 500 MAP antigens detected, 47 were identified as potential candidates for diagnosis of early stages of JD infection (Bannantine *et al.,* 2017; Li *et al.,* 2017). In 2019, this list was updated with the addition of 49 new reactive antigens identified by MAP protein arrays, instead of *Mycobacterium tuberculosis* (Mtb) arrays as in the previous study (Li *et al.,* 2019). Furthermore, Li and others were able to detect antigens specific to different stages of the disease in addition to those expressed throughout the infection (Li *et al.,* 2019). Thus, the use of MAP-specific antigens as a sero-diagnostic tool has been shown to increase the sensitivity and specificity of detection and could potentially also distinguish the stage of the disease.

2.1.3.3 ADDITIONAL CONTROL INTERVENTIONS

Johne's disease requires control at the population level. In addition to early detection and immediate culling of infected animals, farmers need to implement good hygienic practices to prevent MAP exposure of neonates/young animals, such as improving calving area hygiene and management of colostrum/milk feeding to avoid fecal-oral transmission (Whittington *et al.,* 2019).

In sum, no single measure alone is able to reduce infection. Using a nested compartmental (NC) modeling approach, researchers have shown that a combination of test and cull with frequent manure removal was more effective than limiting contacts between calf-adult cow, raising calves in a disease-free herd, or colostrum management (Konboon *et al.,* 2018). However, each farm needs to customize and establish cost-effective best practices within their operations for reducing MAP prevalence. Improving hygiene would not only reduce the prevalence of infectious disease, it will also improve animal health and animal welfare.

2.2. MAP GENOME

Mycobacterium avium subsp. *paratuberculosis* is a member of the *M. avium* complex, which also includes *M. avium* subsp. *avium*, *M. intracellulare*, and *M. silvaticum*, with over 90% similarity at the nucleotide level but major differences in terms of host tropism, microbiological phenotypes, and pathogenicity (Motiwala *et al.,* 2006). In 2005, Li *et al.* described the complete genome of MAP K10, a bovine clinical isolate, providing a strong foundation for investigations into the genetics, evolution, physiology and virulence of MAP. MAP K10 has a single, circular genome of 4,829,781 base pairs, a GC content of 69.3%, and encodes 4,350 open reading frames (ORFs). They also identified 17 copies of the insertion sequence IS*900*, seven copies of IS*1311* and three copies of IS*Mav2* (Li *et al.,* 2005). All these insertion sequences were previously described (Collins *et al.,*1989; Whittington *et al.,*1998; Strommenger *et al.,*2006). However, the defined coordinates of these loci in the genome allowed investigators to use them as markers for genotyping and epidemiological purposes (Dombek *et al.,* 2000). Comparative genome analysis by Li and others identified genes that are likely involved in the pathogenesis of MAP, including multiple genes that differed between MAP and other mycobacterial species and thought to be

involved in virulence, metabolism, and survival (2005). One of the most significant findings was a truncation observed in *mbtA*, a gene known to initiate the synthesis of mycobactins (Quadri *et al.,* 1998; Li *et al.,* 2005). When compared to *Mycobacterium avium* subsp. *avium* (*Mav*) and *Mycobacterium tuberculosis* (*Mtb*)*, mbtA* in MAP is shorter, encoding 337 residues, while *Mav* encoded 565-residue polypeptides and *Mtb* 551-residues (Li *et al.,* 2005).

All these features might be linked to MAP-specific genes and specific genetic elements like the large sequence polymorphisms (LSPs). Mycobacterial genomics studies showed that genome reduction through the loss of LSPs is a major contributor for genetic diversity (Brosch *et al.,* 2002). Numerous LSPs have been identified by a comparative genomic approach between MAP and *Mav*. There are LSPs present in the former but not in the latter and vice versa (Semret *et al.,* 2005). The *MAP* genome includes six specific genomic insertions: LSP^P4, LSP^P11, LSP^P12, LSP^P14, LSP^P15 and LSP^P16; together they comprise ~125kb and 82 ORFs (Alexander *et al.,* 2009). All these genetic elements are known to be essential for MAP survival *in vitro*, increasing MAP's fitness as a obligate pathogen (Wang *et al.,* 2014).

The genomic deletion of LSP^P8 is common to all MAP strains, while other deletions are suggestive for the two major lineages of *MAP*: the sheep strain and the cattle strain (Alexander *et al.,* 2009). These two lineages have been described by extensive molecular strain typing and comparative genomic studies (Bannatine *et al.,* 2012). MAP Type I (sheep) and MAP Type II (cattle) strains were first identified by restriction endonuclease analysis and DNA hybridization from fecal samples or intestinal tissues from cattle, sheep and goats (Collins, Gabric, de Lisle 1990). A Type III strain was originally described as an intermediate strain, but the existence of the third strain was not supported by LSP analysis (Alexander *et al.,* 2009). Due to their high

genetic similarity with Type I isolates, the intermediate strain has since been classified as a sub-lineage of the MAP sheep strain (Castellanos, Domingues, Aranaz 2011).

Identification of LSPs specific for or deleted from MAP provided a better understanding of the evolution of this intriguing pathogenic mycobacteria. Out of the six MAP-specific genomic islands, LSP 14 and 15 are predicted to encode several metal uptake systems. Additionally, Stratmann and others described a putative 38kb MAP-specific ABC transporter operon (*mpt*) locus comprising six ORFs designated *mpt*ABCDEF (2004) with highest similarity to other genera of the order Actinomycetales such as *Salinispora* or *Bifidobacter* and to a lesser extent to other mycobacteria (Meißner *et al.,* 2014). Furthermore, the 38kb MAP-specific operon is immediately preceded by two Fur boxes and located on LSP14 (Stratmann *et al.,* 2004; Alexander *et al.,* 2009). In addition to the *mpt* operon, functional genomics revealed a *fep* cluster encoding a protein involved in the uptake of Fe^{3+} and a *sid* cluster encoding non-ribosome-dependent heterocyclic siderophore synthases (Stratmann *et al.,* 2004). Separated by 4.8kb from LSP^P14, LSP^P15 contains another putative metal transport operon and includes a Fur-like transcriptional regulator (Alexander *et al.,* 2009). In 2016, using transposon mutagenesis, Wang and others reported that LSP^P15 was a horizontally transferred genomic island. Furthermore, with bioinformatics tools and *in vitro* experiments, they demonstrated that LSP^P15 provides an alternative iron uptake system.

The complete genome sequencing of MAP-K10 and other studies that followed provided basic insights on an alternative iron uptake system for MAP survival and virulence, but much remains unknown. Future studies investigating MAP transcriptional responses to iron availability can potentially identify crucial genes involved in pathogenesis that could be considered relevant targets for vaccine development, diagnostics, or improvements to *in vitro* MAP culture.

2.3. IRON IN ALL ORGANISMS

Iron is the fourth most abundant element in the Earth and it is essential for plants, bacteria, animals and humans; it is involved in critical roles like oxygen transportation by haemoglobin, energy production by electron transfer in the mitochondrial respiration process, and DNA synthesis (Lasocki, Gaillard and Rineau, 2014). Furthermore, cells employ iron as a cofactor due to its flexible coordination chemistry and redox reactivity (Wang and Pantopoulos, 2011). Existing in both ferrous (Fe^{2+}) and ferric (Fe^{3+}) oxidation states gives iron the ability to act as a catalyst in various biological systems (Messenger and Barclay, 1983; Frawley and Fang, 2014). However, this catalytic activity can also potentiate iron toxicity, and thus, to prevent intoxication without also inducing iron limitation, iron levels must be tightly controlled (Hood and Skaar, 2012).

In mammals, more than half of the total iron is found in the form of heme, a hydrophobic Fe-based compound that is used as a prosthetic group in proteins to perform essential biochemical processes (Nuñez, Sakamoto and Soares, 2018). Containing the bulk of iron, heme groups known as hemoglobin are a main target for mammalian pathogens, followed by ferritin, transferrin and myoglobin (Ganz and Nemeth, 2015). Although iron is one of the most abundant elements on Earth, upon exposure to oxygen, iron forms insoluble oxides (Fe^{3+}) that are unavailable for absorption (Nuñez, Sakamoto and Soares, 2018). Mammals utilize Fe^{2+} from their diet and through conservation and recycling mechanisms (Nuñez, Sakamoto and Soares, 2018). Iron homeostasis is achieved through uptake and recycling systems such as enterocytes and splenic macrophages, hepatocytes functioning as storage elements, and through functional cooperation of erythroid cells involved with transportation, utilization and storage of iron (Gozzelino and Arosio, 2016; Muckenthaler *et al.,* 2017). The master regulator of iron homeostasis in mammals

is a peptide hormone produced by the liver known as hepcidin (HEP) (Golonka, Yeoh and Vijay-Kumar, 2019). This peptide has microbicidal properties and it is strongly induced during inflammation and infections (Michels et al, 2015). As stated before, iron levels must be tightly controlled, as disruption of its homeostasis has been shown to be responsible in numerous human pathologies. For example, decreased levels of iron and subsequent depletion of its storage result in low supplementation of iron to organs – a condition known as anemia – characterized by a lack of iron within hemoglobin (Muckenthaler *et al.*, 2017). On the other hand, an oversupply of iron is proportional in severity to iron deficiency, and can result in damage to the liver (hemochromatosis), endocrine glands and cardiac myocytes (Michels *et al.,* 2015).

Benefiting from the necessity for fine tuning, mammals limit access to essential metals as iron, limiting infection by pathogenic organisms through a process known as nutritional immunity (Hood and Skaar, 2012). During infection, the host can activate production of the cytokine interleukin (IL)-6 to induce expression of HEP that binds to ferroportin (Fpn), the sole mammalian cellular iron exporter, expressed by macrophages and duodenal epithelial cells (Golonka, Yeoh and Vijay-Kumar, 2019). Ferroportin exports iron into plasma from absorptive enterocytes, from macrophages that recycle the iron of senescent erythrocytes, and from hepatocytes that store iron (Ganz and Nemeth 2012). When HEP-Fpn interaction happens, it induces endocytosis and lysosomal degradation of both molecules, resulting in a rapid reduction of serum iron levels, a consequence of iron being trapped within macrophages (Nemeth *et al.,* 2004; Rodriguez *et al.,* 2014; Rishi, Wallace and Subramaniam, 2015).

Although this mechanism is effective at limiting iron for extracellular pathogens and restricting their growth, studies have shown that it can also be beneficial for intracellular pathogens like *Salmonella* spp., *Chlamydia* spp., *Legionella pneumophila* and *Mycobacterium tuberculosis*

(*Mtb*) (Collins, 2008; Golonka, Yeoh and Vijay-Kumar, 2019). In response to systemic bacterial infection, macrophages themselves produce extra hepcidin, resulting in downregulation of ferroportin, increased levels of iron and intracellular bacterial growth (Collins, 2008; Paradkar *et al.,* 2008). Pathogens require 10^{11}-10^{12}-fold higher concentrations of free iron than what is physiologically available within hosts (Andrews, Robinson and Rodríguez-Quiñones, 2003). Hosts utilize iron-binding proteins like lactoferrin, transferrin and siderocalin to deplete free-iron and supply the fuel for reactive oxygen species (ROS) generation, which enhances bactericidal activity through oxidative killing (Ganz, 2012; Golonka, Yeoh and Vijay-Kumar, 2019). To subvert host nutritional immunity, bacteria express specialized iron acquisition systems using iron-binding molecules known as siderophores (Golonka, Yeoh and Vijay-Kumar, 2019). Siderophores are iron chelators produced by bacteria, fungi, and plants. For bacteria, siderophores scavenge iron from the host and return it back to the pathogen. They are divided into three categories: catecholate, trihydroxamate and mixed siderophores, and the differences between them are associated with bacterial virulence (Golonka, Yeoh and Vijay-Kumar, 2019). In mycobacteria they known as are mycobactins and exochelins, in *Listeria monocytogenes* they are TF-like, and in most Gram-negative bacteria called ferric enterobactins (Schaible and Kayfmann, 2004).

In the battle between host and pathogen, the immune system secretes lipocalins to sequester cateocholate-type siderophores (Lcn2), blocking iron availability for bacteria (Flo *et al.,* 2004). In mammals, neutrophils as well as the liver and gut epithelia are the main sources of Lcn2, and although it is constitutively expressed throughout the body, its levels are increased during host nutritional immunity as an antimicrobial mechanism (Bachman, Miller and Weiser, 2009). Consequently, several pathogens such as *K. pneumoniae, S. aureus, S. typhimurium* and *E. coli*

express Lcn2-resistant or stealth siderophores (Bachman *et al.,* 2012; Lu *et al.,* 2019). Thus, the interaction of bacterial siderophores and Lcn2 is one of the factors that can determine pathogen virulence.

In summary, the pathogen-host interplay and competition for cellular iron requires coordination of multiple biological processes and where both sides require iron homeostasis for survival.

2.4. IRON IN MAP

Iron is essential for life, and MAP is additionally well-known for its unique iron requirements for *in vitro* growth. Unlike other pathogenic mycobacteria, MAP requires ferric mycobactin in order to grow in culture media (Lepper and Wilks, 1988). Mycobactins are siderophores, iron-binding compounds produced by mycobacteria that are responsible for the transportation of ferric iron from host to bacteria (Barclay, Ewing and Ratledge, 1985). Most mycobacteria produce siderophores containing two salicylates divided into subgroups based on lipid solubility: lipophilic and hydrophilic (LaMarca *et al.,* 2004). The former is the cell-associated form of mycobactin, while the latter – also known as carboxymycobactin – is excreted from the cell (Byers, 2013). Studies in *M. tuberculosis* showed that a group of 10 genes, *mbtA-mbtJ,* are responsible for the production of mycobactin and carboxymycobactins to aid in transportation of iron into bacterial cells. *In vitro* assays showed their expression upregulated in response to low iron conditions (McMahon, Rush and Thomas, 2012; Reddy *et al.,* 2013). A comparative genomic analysis of the MAP K10 genome showed a truncation in the *mbtA* gene. This 150bp truncation in *mbtA* was suggested as the likely cause for the *in vitro* mycobactin dependency of MAP (Li *et al.,* 2005). However, studies showed that while the MAP sheep strain does not have an *mbtA* truncation, it is still unable to synthesize mycobactin (Wang *et al.,* 2016). From

previous studies in our lab, we know that *mbt* operon promoter is active and genes are transcribed by MAP inside macrophages and in tissues of infected animals, but it's still unknown what causes mycobactin dependence *in vitro* (Janagama *et al.,* 2009; Janagama *et al.,* 2010).

2.4.1. IRON-DEPENDENT REGULATOR

In MAP, the *mbt* operon is directly involved in iron metabolism and regulated by the iron-dependent regulator protein (IdeR) (Janagama *et al.,* 2009). IdeR is a global regulator of iron response and belongs to the diphtheria repressor (DtxR) family of transcription regulators. In the presence of iron, IdeR binds to a 19bp promoter sequence known as "iron box," repressing transcription of iron acquisition genes and activating genes related to iron-storage (Rodriguez *et al.,* 2002). In *Mtb,* genes regulated by IdeR were putative transporters, proteins involved in siderophore synthesis and iron storage, members of the PE/PPE family, a membrane protein involved in virulence, transcriptional regulators, and enzymes involved in lipid metabolism (Rodriguez *et al.,* 2002).

IdeR in MAP (*MAP2827*) has 93% amino acid identity to its ortholog in *Mtb*. Using a DNase footprinting assay in combination with a gel shift assay, Janagama and others (2009) confirmed that in the presence of iron, *MAP2827* binds to the 19bp iron box on the MAP genome, repressing transcription of *mbtB*, an iron acquisition protein. On the other hand, analysis of MAP *bfrA*, encoding an iron storage protein, showed differential iron regulation between sheep and cattle strains, with the former presenting a defective iron storage system (Janagama *et al.,* 2009; Janagama *et al.,* 2010). The disruption of *mbtE* in the MAP sheep strain might be explained as adaptation to differences in metabolic pathways used by cattle and sheep (Wang *et al.,* 2016). Janagama *et al*. (2010) also reported differences in metabolic pathways used by the two strains to

adapt to iron limitation condition, where aconitase, succinate dehydrogenase and superoxide dismutase downregulation was observed only in the cattle strain, suggesting that an iron-sparing response, repressing the expression of iron-rich proteins when iron is limiting, was exclusive to cattle strains of MAP.

2.4.2. FERRIC UPTAKE REGULATOR

Adding to the complexity surrounding mycobactin and MAP, multiple investigators have suggested the existence of an alternative iron uptake mechanism (Stratmann *et al.,* 2004; Eckelt *et al.,* 2014; Wang *et al.,* 2016). The mycobacterium paratuberculosis transporter (*mpt*) gene cluster was the main candidate, as it encodes two predicted transporters homologous to the iron uptake transporter (IrtAB) of *Mtb*: *mptABC* and *mptDEF.* (Stratmann *et al.,* 2004). The IrtAB proteins are part of the iron acquisition machinery in *M. tuberculosis*, are involved in transport of iron into the cell, and are required for survival and virulence in iron-deficient conditions (Rodriguez and Smith, 2006). In addition, functional genomic analysis of the *mpt* operon in MAP showed similarities consistent with an involvement in molecular processes involved with iron uptake (Stratmann *et al.,* 2004). These findings, in conjunction with the presence of two putative Fur boxes on the MAP genome, strongly suggest that this operon is involved in an alternative iron uptake mechanism.

The ferric uptake regulator (Fur) protein, first described in *E. coli*, is well-known for controlling iron metabolism, regulating defenses against oxidative stresses in most prokaryotic organisms, and is considered the master regulator of iron homeostasis (Hantke, 1981; Ma, Faulkner and Helmann, 2012). The Fur protein in *E. coli* is a 17kDa polypeptide that recognizes a 19bp consensus sequence known as the “fur box” (GATAATGATwATCATTATC; w = A or T) found

in the promoter of regulated genes, and Fur acts as a transcriptional repressor of iron-regulated promoters (Escolar, Pérez-Martín and Lorenzo, 1999). Fur genes have been well-characterized in many Gram-negative bacteria: *Yersinia, Salmonella, Campylobacter, Legionella, Helicobacter pylori* and also in Gram-positive bacteria like *Bacillus subtilis* and *Staphylococcus*. Additionally, Fur shares its nomenclature with a superfamily of metal sensors: Zur (zinc uptake), Mur (manganese uptake), Nur (nickel uptake), PerR (peroxide stress response), and Irr (heme-dependent iron responsive regulator) (Lee and Helmann, 2007).

To date, there have been characterizations of two Fur family members in *MAP*: FurA and FurB, also known as Per (peroxidase stress response) and Zur (zinc uptake repressor), respectively. Furthermore, the MAP genome contains a predicted Fur-like gene, *MAP3773c* in LSP15, a metal transport MAP-specific operon *(*Alexander *et al.,* 2009; Eckelt *et al.,* 2014; Eckelt *et al.,* 2015). Mycobacterial *furA* is located upstream of *katG, a* catalase-peroxidase gene, and it has been proposed that FurA is involved in mycobacterial oxidative stress survival (Zahrt *et al.,* 2001; Eckelt *et al.,* 2015). Using a MAP *furA* deletion, Eckelt and others (2015) were able to confirm the role of FurA of MAP in intracellular survival and response to oxidative stress. They also showed that FurA provides activation of genes that contribute for MAP long term survival in macrophages.

The MAP genome also contains *furB,* producing a protein responsible for zinc homeostasis in many bacteria. Using *M. smegmatis* as a heterologous model and studying zinc starvation, it was demonstrated that FurB was also responsible for the regulation of 70 MAP genes, with 45 of them present in the LSP14 and LSP15 genomic islands (Eckelt *et al.,* 2014).

Despite numerous studies of metal ion sensing by Fur family proteins in bacteria, not a lot information is known for MAP.

2.5. FINAL CONSIDERATIONS

Mycobacterium avium subsp. *paratuberculosis* (MAP), the causative agent of Johne's disease (JD), imposes significant economic losses around the world to the dairy cattle industry. While culling of JD-infected cattle is necessary to control transmission to other herds, this method is problematic, and when used alone, cannot fully address JD prevention and eradication. *Mycobacterium avium* subsp. *paratuberculosis* is well known for its unique iron requirement for *in vitro* growth, but this observation has not been fully explained. A deeper investigation of the large genomic insertions and deletions specific to the MAP genome might provide additional information on the iron uptake systems. Understanding the role of a novel genetic element will aid in better understanding of a key virulence (and *in vivo* survival) element of MAP.

REFERENCES

REFERENCES

Alexander, D.C., Turenne, C.Y., Behr, M.A. (2009) Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. J Bacteriol. 191(3):1018- 25. DOI: 10.1128/JB.01340-08

Alonso, S., Pethe, K., Russell, D.G. and Purdy, G.E. (2007) Lysosomal killing of *Mycobacterium* mediated by ubiquitin-derived peptides is enhanced by autophagy. *Proceedings of the National Academy of Sciences of the USA* 104, 6031–6036.

Andrews, S.C., Robinson, A.K., Rodríguez-Quiñones, F. (2003). Bacterial iron homeostasis. FEMS Microbiology Reviews. 27, 215-237. DOI:10.1016/S0168-6445(03)00055-X

Arsenault, R.J., Maattanen, P., Daigle, J., Potter, A., Griebel, P., Napper, S. (2014) From mouth to macrophage: mechanisms of innate immune subversion by *Mycobacterium avium* subsp. *paratuberculosis.* BMC Genomics. 45-54 DOI:10.1186/1297-9716-45-54

Australian Pesticides and Veterinary Medicines Authority. Trade advice notice on the product Silirum Vaccine. Department of the Prime Minister, February 2014. APVMA 59642

Bachman, M.A., Miller, V.L., Weiser, J.N. (2009). Mucosal Lipocalin 2 Has Pro-Inflammatory and Iron- Sequestering Effects in Response to Bacterial Enterobactin. Plos Pathogens. doi:10.1371/journal.ppat.1000622

Bachman, M.A., Lenio, S., Schimdt, L., Oyler, J.E., Weiser, J.N. (2012). Interaction of Lipocalin 2, Transferrin, and Siderophores Determines the Replicative Niche of *Klebsiella pneumoniae* during Pneumonia. mBio. DOI:10.1128/mBio.00224-11.

Bannantine, J.P., Baechler, E., Zhang, Q., Li, L., Kapur, V. (2002) Genome scale comparison of *Mycobacterium avium* subsp. *paratuberculosis* Reveals potential diagnostic sequences. J Clin Microbiol. 40(4):1303-10. doi: 10.1128/JCM.40.4.1303–1310.2002

Bannantine, J.P., Wu, C., Hsu, C., Zhou, S., Schwartz, D.C., Bayles, D.O., Paustian, M.L., Alt, D.P., Sreevatsan, S., Kapur, V., Talaat, A.M. (2012). Genome sequencing of ovine isolates of *Mycobacterium avium* subspecies *paratuberculosis* offers insights into host association. BMC Genomics. 13:89. DOI:10.1186/1471-2164-13-89.

Bannantine, J.P., Bermudez, L.E. (2013) No Holes Barred: Invasion of the Intestinal Mucosa by *Mycobacterium avium* subsp. *paratuberculosis.* Infection and Immunity 81, 3960-65. doi:10.1128/IAI.00575-13

Bannantine, J.P., Hinnes II, M.E., Bermudez, L.E., Talaat, A.M., Sreevatsan, S., Stabel, J.R., Change, Y.F., Coussens, P.M., Barletta, R.G., Davis, W.C., Collins, D.M., Grohn, Y.T., Kapur,

V. (2014a). A rational framework for evaluating the next generation of vaccines against *Mycobacterium avium* subspecies *paratuberculosis*. Frontiers in Cellular and Infection Microbiology. DOI:10.3389/fcimb.2014.00126

Bannantine, J.P., Evermann, J.L., Rose, S.J., Babrak, L., Katani, R., Barletta, R.G., Talaat, A.M., Grohn, Y.T., Chang, Y., Kapur, V., Bermudez, L.E. (2014b). Frontiers in Cellular and Infection Microbiology. Doi:10.3389/fcimb.2014.00088

Bannantine, J.P., Campo, J.J., Li, L., Randall, A., Pablo, J., Praul, C.A., Garay, J.A.R., Stabel, J.R., Kapur, V. (2017). Identifcation of Novel Seroreactive Antigens in Johne's Disease Cattle by Using the Mycobacterium Tuberculosis Protein Array. Clin Vaccine Immunol. 24(7). doi:10.1128/CVI.00081-17

Barclay, R., Ratledge, C. (1988) Mycobactins and exochelins of *Mycobacterium tuberculosis, M. bovis, Mycobacterium africanum* and other related species. J Gen Microbiol 134,771–776

Barkema, H.W., Orsel, K., Nielsen, S.S., Koets, A.P., Rutten, V.P.M.G., Bannantine, J.P. (2017). Knowlodge gaps that hamper prevention and control of *Mycobacterium avium* subspecies *paratuberculosis* infection. Transbounnd Emerg Diseases 65 (1):125-149. DOI:10.1111/tbed.12723

Bermudez, L.E., Petrofsky M, Sommer S, Barletta R.G. (2010). Peyer's patch-deficient mice demonstrate that *Mycobacterium avium* subsp. *paratuberculosis* translocates across the mucosal barrier via both M cells and enterocytes but has inefficient dissemination. Infect Immun. 78:3570-3577. DOI: 10.1128/IAI.01411-09.

Brennan, M.J. (2017). The Enigmatic PE/PPE Multigene Family of Mycobacteria and Tuberculosis Vaccination. Infect Immun. DOI:10.1128/IAI.00969-16

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc Natl Acad Sci USA. 99, 3684-9.

Brumell J.H., Scidmore, M.A. (2007). Manipulation of rab GTPase function by intracellular bacterial pathogens. Microbiol Mol Biol Rev. (4):636-52. Doi: 10.1128/MMBR.00023-07.

Byers, R.B. Iron Acquisition by the genus Mycobacterium. History, Mechanisms, Role of Siderocalin, anti-tuberculosis drug development. Springer, 2013.

Castellanos, E., Juan, L., Domínguez, L., Aranaz, A. (2011) Progress in molecular typing of *Mycobacterium avium* subspecies *paratuberculosis*. Res Vet Sci DOI:10.1016/j.rvsc.2011.05.017.

Chandrangsu, P., Rensing, C., Helmann, J. (2017) Metal homeostasis and resistance in bacteria. Nat Rev Microbiol. 15(6):338-350. doi: 10.1038/nrmicro.2017.15.

Cheville N.F., Hostetter J., Thomsen B.V., Simutis F., Vanloubbeeck Y., Steadham E. (2001). Intracellular trafficking of *Mycobacterium avium* ss. Paratuberculosis in macrophages. Dtsch Tierarztl Wochenschr. 108:236–243

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S. & other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393, 537–544.

Collins, D.M., Gabric, D.M., De Lisle, G.W. (1989). Identification of a repetitive DNA sequence specific to *Mycobacterium paratuberculosis.* FEMS Microbiology Letters. DOI:10.1111/j.1574-6968.1989.tb03440.x

Collins, D.M., Gabric, D.M., de Leslie, G.W. (1990). Identification of two groups of *Mycobacterium paratuberculosis* strains by restriction endonuclease analysis and DNA hybridization. Journal of Clinical Microbiology. 28, 1591-96.

Collins, H.L. (2008). Withholding iron as a cellular defence mechanism- friend or foe? European Journal of Immunology. 38, 1803-06. DOI:10.1002/eji.200838505

Coussens, P. (2001) *M. avium* subspecies *paratuberculosis* and the bovine immune system. *Animal Health Research Reviews* 2, 141–161.

Coussens, P., DeKuiper, J.L., Shoyama, F.M., Brenner, E., Lamont, E.A., Kabara, E., Sreevatsan, S. (2020). Host-Pathogen Interactions and Intracellular Survival of *Mycobacterium avium* subsp. *paratuberculosis.* In M.A.Behr (Ed.), K.S., V.K. *Paratuberculosis – Organism, Disease, Control* - 2nd Edition (pp 120-139). Oxfordshire (UK): CABI

Danelishvili, L., Yamazaki, Y., Selker, J., and Bermudez, L. E. (2010). Secreted Mycobacterium tuberculosis Rv3654c and Rv3655c proteins participate in the suppression of macrophage apoptosis. PLoS ONE. doi:10.1371/journal.pone.0010474

Dombek, P.E., Johnson, L.K., Zimmerley, S.T., Sadowsky, M.J. (2000). Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl Environ Microbiol. 66, 2572-7. DOI:10.1128/aem.66.6.2572-2577.2000

Eckelt, E., Jarek, M., Fromke, C., Meens, J., Goethe, R. (2014) Identification of a lineage specific zinc responsive genomic island in *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics. 15:1076. doi: 10.1186/1471-2164-15-1076

Eckelt, E., Meißner, T., Meens, J., Laarmann, K., Nerlich, A., Jarek, M., Siegfried Weiss, S., Gerlach, G-F., Goethe, R. (2015). FurA contributes to the oxidative stress response regulation of *Mycobacterium avium* ssp. *paratuberculosis.* Frontiers in Microbiology. DOI:10.3389/fmicb.2015.00016.

Escolar, L., Pérez-Martín, J., de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. J. Bacteriol. 181, 6223-9.

Everman, J.L., Danelishvili, L., Flores, L.G., Bermudez, L.E. (2018). MAP1203 promotes *Mycobacterium avium* subsp. *paratuberculosis* binding and invasion to bovine epithelial cells. Front. Cell. Infect. Microbiol. Doi: 10.3389/fcimb.2018.00217

Flo, T.H., Smith, K.D., Sato, S., Rodriguez, D.J., Holmes, M.A., Strong, R.K., Akira, S., Aderem, A. (2004). Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. Letters to Nature. 432, 917-921. DOI:10.1038/nature03104.

Foddai, A., Elliott, C.T., Grant, I.R. (2009) Optimization of a Phage Amplification Assay to Permit accurate enumeration of Viable *Mycobacterium avium* subsp. *paratuberculosis* cells. Applied and Environmental Microbiology. 75, 3896-3902. DOI:10.1128/aem.00294-09

Frawley, E.R., Fang, F.C. (2014). The ins and outs of bacterial iron metabolism. Mol. Microbiol 93, 609-16. DOI:10.1111/mmi.12709

Ganz, T. (2012). Macrophages and Systemic iron homeostasis. Journal of Innate Immunity. 4, 446-453. DOI:10.1159/000336423

Ganz, T., Nemeth, E. (2015). Iron homeostasis in host defence and inflammation. Nat. Rev. Immunol. 15, 500-10. DOI:10.1038/nri3863.

Garcia, A.B., Shalloo, L. (2015) The economic impact and control of paratuberculosis in cattle. J. Dairy Sci. 98:5019–5039. doi:10.3168/jds.2014-9241

Golonka, R., Yeoh, B.S., Vijay-Kumar, M. (2019). The iron Tug-of-War between bacterial siderophores and Innate Immunity. J Innate immune. 11, 249-262. DOI:10.1159/000494627

Good, M., Duignan, A. (2011). Perspectives on the history of bovine TB and the role of tuberculin in bovine TB eradication. Vet. Med. Intl. 2011:1–11

Gozzelino, R., Arosio, P. (2016). Iron homeostasis in Health and Disease. Int J Mol Sci. 17:1. DOI: 10.3390/ijms17010130.

Hantke, K. (1981). Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol Gen Genet. 182, 288-92. DOI:10.1007/bf00269672

Hantke, K. (2001) Iron and metal regulation in bacteria. Current Opinion in Microbiology. 4:172-177.

He, Z., De Buck, J. (2010). Localization of proteins in the cell wall of Mycobacterium avium subsp. paratuberculosis K10 by proteomic analysis. Proteome Science. 8:21. Doi:

Hood, M.I., Skaar, E.P. (2012). Nutritional immunity: transition metals at the pathogen–host interface. Nature. 10, 525-537. DOI: 10.1038/nrmicro2836

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Janagama, H.K., Lamont, E.A., George, S., Bannantine, J.P., Xu, W.W., Tu, Z.J., Wells, S.J., Schefers, J., Sreevatsan, S. (2010) Primary transcriptomes of *Mycobacterium avium* subsp. *paratuberculosis* reveal proprietary pathways in tissue and macrophages. BMC Genomics. 11, 561. doi: 10.1186/1471-2164-11-561.

Kabara, E., Coussens, P.M. (2012). Infection of primary bovine macrophages with *Mycobacterium avium* subsp. *paratuberculosis* suppresses host cell apoptosis. Frontiers in Micro. Doi:10.3389/fmicb.2012.00215.

Kennedy, D., Benedictus, G., Nielsen, S., Lybeck, K., Schwan, E., Frössling, J., Sergeant, E., Kelton, D., Nauholz, H. (2017). Guidelines for certification with respect to the movement of livestock for *Mycobacterium avium* subsp *paratuberculosis* (MAP) infection V1.3. Paratuberculosis News. 4:3–17.

Konboon, M., Bani-Yaghoub, M., Pithua, P.O., Rhee, N., Aly, S.S. (2018). A nested comparmental model to assess the efficacy of paratuberculosis control measures on U.S. dairy farms. Plos One. DOI:10.1371/journal.pone.0203190

LaMarca, B.B.D., Zhu, W., Arceneaux, J.E.L., Byers, B.R., Lundrigan, M.D. (2004). Participation of *fad* and *mbt* Genes in Synthesis of Mycobactin in *Mycobacterium smegmatis.* Journal of Bacteriology. 186, 374-382. DOI:10.1128/JB.186.2.374-382.2004.

Lamont, E.A., Xu, W.W., Sreevatsan, S. (2013) Host-*Mycobacterium avium* subsp. *paratuberculosis* interactome reveals a novel iron assimilation mechanism linked to nitric oxide stress during early infection. BMC Genomics. 2164 DOI:10.1186/1471-2164-14-694

Lamont, E. A., Talaat, A. M., Coussens, P. M., Bannantine, J. P., Grohn, Y. T., Katani, R., et al. (2014). Screening of Mycobacterium avium subspecies *paratuberculosis* mutants as vaccine candidates in a primary bovine monocyte-derived macrophage model. Front. Cell. Infect. Microbiol. 4:87. doi: 10.3389/fcimb.2014.00087

Lasocki, S., Gaillard, T., Rineau, E. (2014). Iron is essential for living! Crit Care. 18:678. DOI:10.1186/s13054-014-0678-7.

Lei L., Hostetter, J.M. (2007). Limited phenotypic and functional maturation of bovine monocyte-derived dendritic cells following *Mycobacterium avium* subspecies *paratuberculosis* infection in vitro. Vet. Immunol. Immunopathol. 120:177–186.

Magombedze, G., Shiri, T., Eda, S., Stabel, J.R. (2017). Inferring biomarkers of *Mycobacterium avium* subspecies *paratuberculosis* infection and disease progression in cattle using experimental data. Scientific Reports. 7: 44765. Doi:10.1038/srep44765.

Ponnusamy S., Periasamy, S., Tripathi, B.N., Pal, A. (2013). *Mycobacterium avium* subsp. *paratuberculosis* invades through M cells and enterocytes across ileal and jejunal mucosa of lambs. Res. Vet. Sci. 94:306-12. DOI: 10.1016/j.rvsc.2012.09.023.

Rowe, M.T. and Grant, I.R. (2006) Mycobacterium avium ssp. paratuberculosis and its potential survival tactics. Letters in Applied Microbiology 42, 305–311.

Secott, T. E., T. L. Lin, and C. C. Wu. (2001). Fibronectin attachment protein homologue mediates fibronectin binding by *Mycobacterium avium* subsp. *paratuberculosis.* Infect. Immun. 69:2075–2082

Secott, T.E., Lin, T.L., Wu, C.C. (2004). *Mycobacterium avium* subsp. *paratuberculosis* fibronectin attachment protein facilitates M-cell targeting and invasion through a fibronectin bridge with host integrins. Infection and Immunity. 72 (7) 3724-3732. Doi: 10.1128/IAI.72.7.3724–3732.2004

Souza, C., Davis, W.C., Eckstein, T.M., Sreevatsan, S., Weiss, D.J. (2013). Mannosylated lipoarabinomannans from *Mycobacterium avium* subsp. *paratuberculosis* alters the inflammatory response by bovine macrophages and suppresses killing of *Mycobacterium avium* subsp. *avium* organisms. PLoS One. doi: 10.1371/journal.pone.0075924

Waddell, L., Rajic, A., Stark, K., McEwen, S.A. (2016). *Mycobacterium avium* ssp. *paratuberculosis* detection in animals, food, water and other sources or vehicles of human exposure: a scoping review of the existing evidence. Prev Vet Med.132:32–48.

Weiss DJ, Evanson OA, de Souza C, Abrahamsen MS. A critical role of interleukin10 in the response of bovine macrophages to infection by *Mycobacterium avium* subsp. *paratuberculosis*. Am J Vet Res. 2005. 66:721-6.

Whittington, R.J. (2009). Factors Affecting Isolation and Identification of *Mycobacterium avium* subsp. *paratuberculosis* from Fecal and Tissue Samples in a Liquid Culture System. Journal of Clin. Micro. 612-622. Doi:10.1128/jcm.01986-08.

Woo, S.R., Heintz, J.A., Albrecht, R., Barletta, R.G. and Czuprynski, C.J. (2007) Life and death in bovine monocytes: the fate of *Mycobacterium avium* subsp. *paratuberculosis.* Microbial Pathogenesis 43, 106–113.

CHAPTER THREE: CHARACTERIZATION OF A *MAP*-SPECIFIC GENOMIC ISLAND USING A *MAP3776C::Tn* INSERTIONAL MUTANT

3.1 ABSTRACT

One of the unique characteristics of *Mycobacterium avium* subsp. *paratuberculosis* (MAP), the causative agent of Johne's disease (JD), is that it requires supplementation of a siderophore, mycobactin, for optimal growth in laboratory media – a functional characteristic that has been implicated for its fastidious nature and slow growth in culture media. Prior studies have characterized the function of a *mycobacterium*-specific transcriptional regulator, IdeR, in iron regulation by MAP strains. A second ferric uptake regulator (Fur)-like element (*MAP3773c*), discovered on a MAP-specific genomic island (LSP15), is upregulated *in vivo* in intestinal lesions. This gene is present is a 5 gene cluster (*MAP3776c-MAP3772c*) that has been proposed to function as an operon. The presence of a second MAP-specific iron regulator is intriguing and needs further functional elucidation. The objective of this study was to test whether the genes on LSP15 (specifically *MAP3773c)* provide an alternative iron regulatory phenotype in MAP. Multiple sequence alignments of well-studied Fur proteins (*E. coli* and *S. typhimurium*) against *MAP3773c* revealed significant amino acid similarity and conserved "Fur" domains, suggesting that it encoded a ferric uptake regulator. Relative gene expression of LSP15 in epithelial cells alone, macrophages alone, or in co-culture revealed an up regulation of *MAP3773c* in epithelial cells, alone or in co-culture, suggesting a possible role for *MAP3773c* in cellular invasion and subsequent survival. A loss of Fur transcription was demonstrated using a transposon mutant of the *MAP3776c* (*MAP3776c*::Tn), a gene, upstream of the Fur-like element (*MAP3773c*) located in the same putative operon likely leading to dysfunction of all downstream genes. Infection of epithelial cell (MAC-T) and its co-culture with monocyte-derived macrophages (MDMs),

revealed that the transposon mutant was invasion deficient and was depleted within 48 hr after cell entry. While this data suggests a likely phenotype for LSP15 genes in *in vivo* and *ex vivo* cell entry and survival suggesting a function in epithelial cell invasion of this organism, further functional analysis with complementation of *MAP3776c::Tn* strain would be necessary to confirm these findings, as transposon mutants are known for their polar effect.

3.2 INTRODUCTION

Johne's disease (JD), is a chronic enteritis of ruminants caused by infection with *Mycobacterium avium* subsp. *paratuberculosis* (MAP). MAP is a gram-positive, acid-fast bacillus, with a complex impermeable waxy cell wall composed of 60% lipids, including mycolic acids (Lombard 2011; Marino *et al.,* 2017). Transmission of JD primarily occurs by the fecal-oral route through the ingestion of contaminated feces, water, colostrum, or milk. Infection usually occurs within the first months of life of the animal and remains sub-clinical for 2-5 years. While in sub-clinical stages, infected animal can shed 10^3 colonies/g of feces to millions or billions of bacteria to the environment (Whittington *et al.,* 1998). JD has a significant economic impact in the dairy herds worldwide due to increased premature culling, replacement costs, decrease of milk production, reduced slaughter values, and increased susceptibility to other disease (Garcia and Shalloo, 2015).

Control of the disease is difficult, as in addition to the lack of sensitivity and specificity of the diagnostic assays, currently available vaccines do not have a high efficacy that is necessary to control JD (Bannantine *et al.,* 2014). Thus, in order to improve control strategies, there is a need to better understand the biology of MAP. The most distinctive phenotype of MAP is its inability to produce detectable mycobactin *in vivo* and *in vitro*, requiring mycobactin, a siderophore, in

laboratory cultivation medium for optimal growth. Initially, it was thought that the truncation on the MAP K10 genome of *mbtA* was the likely cause of the mycobactin dependence *in vitro* (Li *et al.,* 2005). However, even without this truncation, sheep strain of MAP is unable to produce mycobactin, furthermore, additional studies have established that *mbt* operon promoter is active and these genes are transcribed by MAP in macrophages and in tissues of infected animals (Zhu *et al.,* 2008; Janagama *et al.,* 2009; Janagama *et al.,* 2010; Wang *et al.*, 2016).

Since the complete genome sequencing of MAP K10, functional characterization of Iron dependent repressor protein (IdeR) in MAP demonstrated that IdeR not only controls mycobactin synthesis but also regulates transcription of genes involved in iron acquisition (*mbtB*), iron storage (*bfrA*) and oxidative stress (Janagama *et al.,* 2009). Furthermore, two MAP-specific genomic islands, large sequence polymorphisms (LSPs) 14 and 15, encode several metal uptake systems (Stratmann *et al.,* 2004; Alexander *et al.,* 2009). In addition, LSPs 14 and 15 carry two ferric uptake regulator (fur) boxes and a Fur-like metal-dependent transcriptional regulator (Stratmann *et al.,* 2004; Alexander *et al.,* 2009). Interestingly, a deep sequencing of mRNA from MAP cultures confirmed the involvement of the lineage specific gene loci LSP14 and LSP15 in metal homeostasis (Eckelt *et al.,* 2014). Bioinformatics analysis predicted that LSP15 encoded an ATP-binding cassette (ABC) transporter (*MAP3776c*), a metal uptake regulator (*MAP3773c)* and a gene that may be involved in cobalamin synthesis (*MAP3772c*) (Alexander, Turenne and Behr, 2009). Additionally, a disruption in *MAP3776c* showed 10-fold reduction in bacterial fitness in a mouse model suggesting the likely importance of the genes in the region of LSP15 during infection (Wang *et al.,* 2014). Furthermore, previous work has demonstrated that genes located on LSP15 are differentially regulated in naturally infected bovine tissues, and

interestingly, *MAP3773c*, a ferric uptake regulator (Fur)-like protein, was downregulated in the tissues and upregulated in experimentally infected macrophages *in vitro* (Janagama *et al.,* 2010). Fur is an iron responsive repressor that controls the expression of genes for siderophore biosynthesis and iron transportation (Hantke, 1981). Furthermore, it has been reported that intracellular iron concentrations in many bacteria are under the control of Fur (Guerinot, 1994). During infection, in a metal-depleted environment, Fur allows for efficient acquisition of iron leading to optimal fitness of the pathogen (Lamont, Xu and Sreevatsan, 2013). Supporting this critical role of Fur in virulence, a deletion of Fur in numerous bacterial pathogens most often resulted in partial or complete attenuation within animal models of infection (Ma, Faulkner and Helmann, 2012). In this study we investigated the phenotype of *MAP3776c* using a Tn insertional mutation of *MAP3776c* located on LSP15.

3.3 MATERIALS AND METHODS

3.3.1 BACTERIAL CULTURE AND CONDITIONS

Mycobacterium avium subsp. *paratuberculosis* K10 and transposon insertion mutant *MAP3776c (MAP3776c*::Tn) strains were grown at 37°C shaking in Middlebrook 7H9 supplemented with 10% OADC (oleic acid, dextrose, catalase) enrichment medium (Thermo Fisher Scientific, Waltham, MA, United States), 0.05% Tween 80, and 2 mg of ferric mycobactin J (Allied Monitor Inc., Fayette, MO, United States) per liter. Antibiotics (µg/ml: kanamycin, 50; hygromycin, 100) were added when necessary. Iron-depleted with Tween 80 medium was prepared using Chelex 100 resin (Bio-Rad) (10 g/liter). *E. coli* TOP10F cells (Invitrogen, Carlsbad, CA, United States) were grown in LB medium 37°C with shaking at 200 RPM.

3.3.2 CONSTRUCTION OF MAP COMPLEMENTATION OF THE TRANSPOSON MUTANT STRAIN

Complementation of *MAP3776c* function of *MAP3776c::Tn* was accomplished by cloning and transformation of the mutant with *MAP3776c. MAP3776c* gene was amplified from MAP K10 DNA, using primers 20 pmoles each primer (fwd: GTGAACCTGATCGCCAAGAT, rev: CTCGGTCTGCGTGTTGTAGA). PCR was carried out by standard methods. Briefly, 25 ng MAP K10 genomic DNA, HotStarTaq Master Mix (Qiagen, Germantown, MD, United States). PCR product was purified using PCR purification kit (Qiagen, Germantown, MD, United States). Sanger sequencing to confirm *MAP3776c* PCR product was performed at the University of Minnesota – Genomics Center.

The constitutive plasmid pSM417 with Hsp60 promoter was used as a vector (Kugadas et *al*., 2016). Double digestion of *MAP3776c* and psM417 was performed using enzymes Pst-1 HF and BamHI-HF (New England Biolabs, Ipswich, MA, United States) followed by Quick ligation protocol (New England BioLabs). The ligation product *MAP3776c*-psM417 was added into the chemically competent *E. coli* TOP10 (Invitrogen, Carlsbald, CA) and single colonies were selected for Sanger sequencing.

Electroporation was carried out following protocol described by Goude & Parish, 2008. A total of 0.5 µg purified *MAP3776c*-psM417 DNA was added into 0.2 mL of electro-competent *MAP3776c::Tn* cells for electroporation and subject to a single pulse of 2.5 kV, 25 µF, with the pulse-controller resistance set at 1000 (ohms) resistance. Cells were recovered with 10 mL of Middlebrook 7H9 supplemented with 10% OADC (oleic acid, dextrose, catalase) enrichment medium (Thermo Fisher Scientific, Waltham, MA, United States) (M7H9-OADC) and after 16 h

incubation at 37° C, cells were plated into solid medium M7H9-OADC containing kanamycin (50 μg/ml) and hygromycin (100 μg/ml) until colonies became visible.

Single colonies were transferred to a liquid M7H9-OADC with kanamycin (50 μg/ml) and hygromycin (100 μg/ml) and incubated at 37° C until OD_{600} of ~0.5 was reached. Successful integration of *MAP3776c* was confirmed by PCR and sequencing. The transposon mutant was kindly provided by Dr. Marcel Behr for these studies (Wang *et al.*, 2014).

3.3.3 PREPARATION OF ELECTROCOMPETENT MYCOBACTERIA

To obtain electrocompetent *MAP3776c::Tn*, a single colony was used to inoculate 10 mL of M7H9-OADC medium containing kanamycin (50 μg/ml) and hygromycin (100 μg/ml) and incubated at 37° C until OD_{600} of ~0.5 was reached. Culture was combined with 100 mL of M7H9-OADC medium and antibiotics until OD_{600} of ~0.5 was reached. The day before cells were harvested 2 M glycine was added. The following day, cells were washed with 10 % glycerol with gradually decreasing volumes (50, 25, 10, 5 mL) and aliquoted into 220 μL and frozen at -80°C. Procedure was carried out following protocol described by Goude & Parish, 2008.

3.3.4 EXTRACTION OF RNA AND QPCR

RNA extraction of each strain – MAP K10, *MAP3776c::Tn* and complemented strains, was performed as described (Kugadas et *al*., 2016). In brief, one milliliter of bacterial culture was collected to a 2 ml sterile screw-cap microcentrifuge tube and 2 ml of TRIzol reagent (Life Technologies, Carlsbad, CA, United States) and 0.6 ml of sterile RNase-free 0.1 mm zirconium beads (BioSpect Products, Burtlesville, OK, United States) were added to tube. Cells were

disrupted in MagNA Lyser (Roche Diagnostics, Sandhofer, Germany) at 7.000rpm for three pulses of 50 s with 1 min intervals on ice. The supernatant was transferred to a 1.5 ml sterile RNase-free microcentrifuge tube, followed by purification using Qiagen RNeasy miniprep (Qiagen, Germantown, MD, United States). DNase treatment with Ambion TurboDNA-free kit (Thermo Fisher Scientific, Waltham, MA, United States) was done according to manufacturer's instructions. cDNA synthesis was performed using the Superscript III First Strand Synthesis SuperMix for qRT PCR (Invitrogen, Carlsbad, CA, United States).

qPCR was conducted using LightCycler 480 II (Roche, Indianapolis, IN, United States) at a volume of 25 µL (5 µL extracted cDNA in 20 µL PCR master-mix). The PCR master-mix consisted of LightCycler 480 SYBR Green 1 Master Mix (Roche, Indianapolis, IN, United States), 0.2mM of each primer.

3.3.5 MAC-T CELL INVASION ASSAY

To determine virulence phenotype of the *MAP3776c::Tn* strain (expected knocked down *MAP3773c* due to an upstream insertional deletion in the putative operon), we performed cell invasion assays with bovine mammary epithelial cells (MAC-T cells). MAC-T cells were kindly provided by Dr Luiz E. Bermudez. They were used in this experiment as this cell-line is considered a surrogate for intestinal epithelial and the source of the cell line, the mammary, is hypothesized to serve as a reservoir for MAP *in vivo* (Patel *et al.,* 2006). Approximately $2.0x10^4$ MAC-T cells were seeded into three T25 flasks (3 flasks per time point) and incubated for 4 days in DMEM containing 10% FBS at 37°C in a humidified chamber at 5% CO_2. Once semiconfluency was reached, MAP was added to the well to allow infection for 3hr at 37°C in a humidified chamber at 5% CO_2 followed by three washings with 1x D-PBS to remove any

nonadherent bacteria. The cells were incubated at 37°C in a humidified chamber for up to 48 hours and further processed for CFU count. CFUs were obtained by Flow Cytometry using the LIVE/DEAD BacLight Bacterial Viability and Counting kit (Thermo Fisher Scientific, Waltham, MA, United States).

3.3.6 CO-CULTURE ASSAY INVASION ASSAY

Assay was conducted using method developed by Lamont, Xu and Sreevatsan (2013). Approximately 2.0×10^4 MAC-T cells were seeded on the apical side of 3.0 μm pore size transwell (12 well plate format) (Corning, Lowell, MA, United States) and incubated in DMEM with 10% FBS for 4 days at 37°C. Once MAC-T cells were 70% confluent, bovine MDMs were seeded at ~ 2.0×10^4 cells in the basolateral chamber and allowed to adhere for 2 h at 37°C. Upon completion of incubation, cells were gently washed 3× using 1X D-PBS to remove nonadherent cells and DMEM medium containing 10% FBS was replaced. MAC-T/bovine MDMs were incubated for additional 2 d. MAC-T cells were infected for 30 min with MAP K10 (OD_{600} = 1.0 and MOI of 10:1). All cells were further washed as before and transwells were removed from the supports and transferred to a new sterile 12 well plate where cells were lysed and transferred to a new microcentrifuge tube for RNA extraction.

3.3.7 SAMPLE PROCESSING

RNA was extracted from two biological replication, each contained twelve samples composed of 4 uninfected host cell types (Macrophage and MAC-T under single cell type or co-cultured conditions), 4 infected host cell types (Macrophage and MAC-T under single cell type or co-cultured conditions), 2 pathogen samples from macrophages and MAC-T cells under single cell

type culture, and 2 pathogen samples from co-cultured macrophages and MAC-T cell. Followed by extraction, RNA was submitted to enrichment and amplification for MAP transcripts using MICROBEnrich (Ambion, Austin, TX) and MessageAmpII bacteria kit (Ambion, Austin, TX) per manufacturer's instructions. Successful elimination of host RNA was determined by the RNA 6000 LabChip kit and Agilent Bioanalyzer (Caliper Technologies Corp., Hopkinton, MA and Agilent Technologies, Santa Clara, CA). RNA integrity and purity were analyzed by RiboGreen assay per manufacturer's instructions (Invitrogen, Carlsbad, CA) and the Agilent Nanochip (Agilent Technologies, Santa Clara, CA). All RNA samples with RNA Integrity Number (RIN) of 5 or greater were submitted for RNA-seq.

3.3.8 RNA-SEQUENCING

A total of 24 samples were submitted for sequencing. The RNA-Seq library was created using the mRNA Seq library preparation kit per manufacturer's instructions (Illumina Inc., San Diego, CA). The purified library was later validated and quantified using the Agilent High Sensitivity Chip (Agilent Technologies, Santa Clara, CA), picogreen assay (Invitrogen, Carlsbad, CA), and KAPA qPCR (KAPA Biosystems, Woburn, MA) as described by the corresponding manufacturer. The Illumina cBOT (Illumina Inc., San Diego, CA) was used for cluster generation. Briefly, the template (samples in 8-plex) was immobilized to a random oligo lawn on the flow cell surface, which was later amplified, linearized, blocked and hybridized to the sequencing primer. The clustered flow cell was then transferred and loaded into the Illumina Genome Analyzer IIx (GaIIx) (Illumina Inc., San Diego, CA). Sequencing was conducted using pair-end chemistry. Twenty million reads (200 bp insert size) and 7.5 million reads (150 bp insert size) were recorded for each bovine and MAP samples, respectively.

3.3.9 RNA-SEQ ANALYSIS

Computational analysis was carried out on the Galaxy platform at the Minnesota Supercomputing Institute at the University of Minnesota. Bovine FASTQ files were mapped to the *B. taurus* genome (Btau 4.0) using TopHat (v2.02) with 2 mismatch setting. The differential gene expression (DGE) was analyzed using Cufflinks (Cuffdiff program). MAP FASTQ files were mapped to the MAP K10 genome using Bowtie with 2 mismatch and the DGE was examined using edgeR program in BioConductor. DGE was determined by a q-value cutoff of 0.05 or p-value cutoff of 0.05 as determined by Cuffdiff and edgeR.

3.4 RESULTS

3.4.1 COMPUTATIONAL CHARACTERIZATION

The sequence alignment produced by Muscle, showed that when compared to the well-studied Fur proteins of other bacteria, *MAP3773c* carried several missense mutations, however the biochemical properties of altered amino acids were similar (Fig. 1). In addition, NCBI database showed that *MAP3773c* has conserved most of the functions related to the Fur protein with a helix-turn-helix motif and a metal binding and a structural Zn (II) binding site (Fig. 2).

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)

MAP_3773c              MSSPAAPRRR---RATVKQRTVLEVLRAQENFR-SAQQLYQDIRQNQQLRIGLTSVYRIL
F.tularensis_Fur       MNSKNLDLKEFGFKVTQPRVEILKLFEKNKDKHLSPDDVFSKLKAQGS-TTGIATVYRVL
K.pneumoniae_Fur       MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLIDMGE-EIGLATVYRVL
S.typhimurium_Fur      MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLIDMGE-EIGLATVYRVL
E.coli_Fur             MTDNNTALKKAGLKVTLPRLKILEVLQEPDNHHVSAEDLYKRLIDMGE-EIGLATVYRVL
E.cloacae_Fur          MTDNNTALKKAGLKVTLPRLKILEVLQGPDNHHVSAEDLYKRLIDMGE-EIGLATVYRVL
                       *..     .    ..*  .  :*:::   .: . *.:::::. :    .   *:::***:*
```

Figure 1. **Sequence alignment generated by Muscle shows that *MAP3773c* is a divergent *fur*.** Asterisk indicates position that has a single, fully conserved residue; colon indicates conservation between groups of strongly similar properties and period indicates conservation between groups of similar properties.

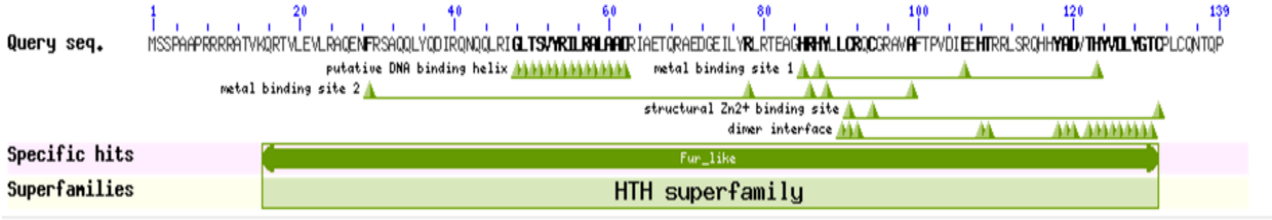


Figure 2. **Conserved domains of *MAP3773c* generated by NCBI**. Figure shows that gene has metal binding site and structural Zn (II) binding site and confirmed to be a Fur-like protein.

3.4.2 COMPLEMENTATION OF *MAP3776C*::TN

For initial investigations of *MAP3773c* function, we used the transposon insertion mutant of *MAP3776c*, located in the MAP-specific genomic island, LSP15 and upstream of *MAP3773c* (Fig. 3). In addition, analysis with PCR showed that disruption of *MAP3776c* leads to a loss of *MAP3773c* transcription in the transposon mutant strain (Fig. 4).

For complementation of the transposon mutant, *MAP3776c* gene from MAP K10 (Fig. 5) was cloned downstream the *hsp60* promoter (*Phsp60*) into the constitutive plasmid psM417 to form *MAP3776c*-psM417 plasmid (Fig 6). Next, *MAP3776c*-psM417 DNA was cloned into *MAP3776c*::Tn and restauration of *MAP3776c* expression was confirmed by PCR (Fig. 7).

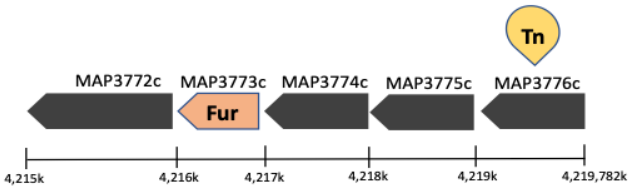


Figure 3. **Representation of MAP-specific genomic island, LSP15, highlighting *MAP3776c* upstream position in relation to *MAP3773c***. Yellow tear drop represents location of transposon mutagenesis.

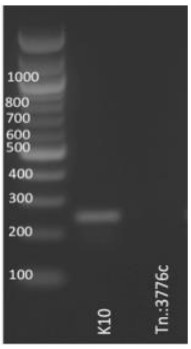


Figure 4. **1 % Agarose gel showing PCR amplification of *MAP3773c***. Lane 1. 1kb ladder; Lane 2. PCR amplification of *MAP3773c* (280bp) from MAP K10 DNA; Lane 3. Absence of PCR amplification of *MAP3773c* from transposon mutant.

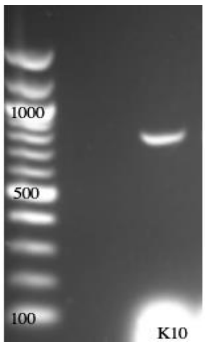


Figure 5. **1 % Agarose gel showing PCR amplification of *MAP3776c***. Lane 1. 1kb ladder; Lane 2. Blank; Lane 3. PCR amplification of *MAP3776c* (800bp) from MAP K10 DNA.

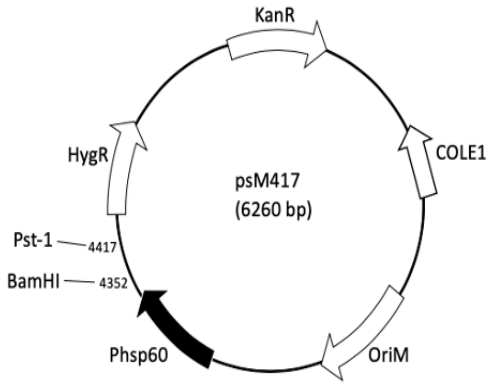


Figure 6. **Representation of psM417 map.** Map indicates location where plasmid was cut (Pst-1 and BamHI) for *MAP3776c* insertion.

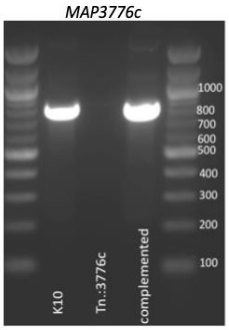


Figure 7. **1 % Agarose gel showing PCR amplification of *MAP3776c* mutant and complemented strains.** Lane 1. 1kb ladder; Lane 2. PCR amplification of *MAP3776c* (800bp) from MAP K10 DNA; Lane 3. Absence of PCR amplification of *MAP3776c* from transposon mutant. Lane 3. PCR amplification of *MAP3776c* (800bp) from complemented strain; Lane 4. 1kb ladder.

3.4.3 *EX VIVO* PHENOTYPE OF THE TRANSPOSON MUTANT OF *MAP3776C*

Epithelial cell (MAC-T cells) invasion assay was carried out to investigate the phenotype of the transposon mutant *MAP3776c.* The efficiencies of invasion and intracellular survival were measured up to 48 h post-infection. Cells were lysed and plated for CFU counts. Comparisons of the mutant against two cattle strains (MAP K10, MAP 1018) and one sheep strain (MAP7565), indicate that the mutant was attenuated in invasion efficiency as well as deficient in survival post invasion, suggesting that the MAP-specific genomic island, LSP15, may play a role virulence of MAP (Fig. 8).

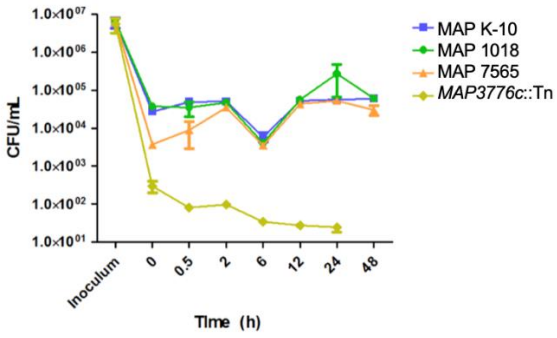
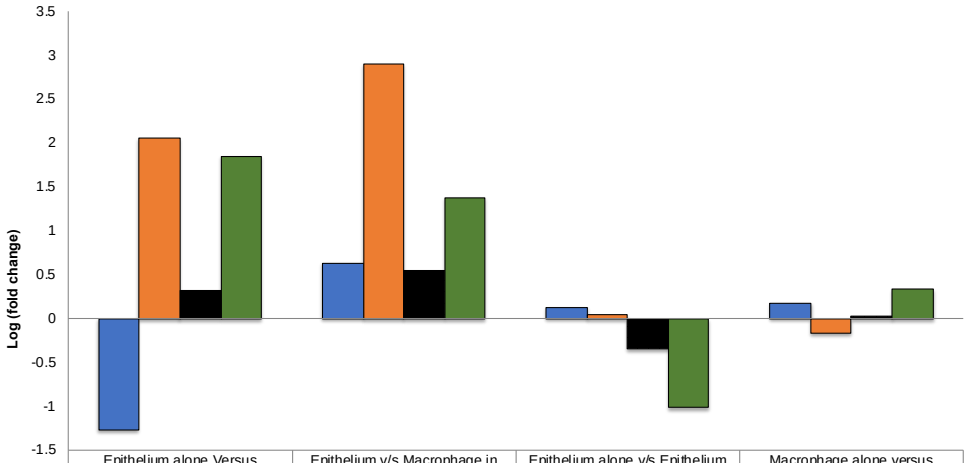


Figure 8. **MAC-T cell invasion comparisons of wild type (MAP K10, MAP 1018 and MAP7565) vs *MAP3776c*::Tn**. Cells were collected at each time pointed indicated in the graphic and plated for CFU count. MAP mutant showed low invasion efficiency throughout the 48 h.

Next, we investigated if *MAP3776c* (and likely *MAP3773c)* was involved in macrophage recruitment during MAP K10 infection in a co-culture system with MAC-T cells and Monocyte-derived macrophages (MDMs), recreating an early infection of MAP. If MAP induced macrophage recruitment, MDMs would permeate the transwell pores into the basolateral chamber. RNA-seq was conducted on MAC-T, macrophages alone, or co-culture, with or without MAP infection. Differential gene expression (DGE) of RNA-seq data from the *in vitro* experiment, revealed that *MAP3773c* gene expression was up regulated ($p < 0.05$) in MAC-T cells either in the presence of macrophages or in epithelium alone when compared to the corresponding uninfected control. The co-culture system provided us the ability to study MAP-host interactions, upregulation of *MAP3773c* in the early stages of infection suggests a possible role of *MAP3773c* in transepithelial migration of MAP (Fig. 9).



	Epithelium alone Versus Macrophage alone	Epithelium v/s Macrophage in co-culture	Epithelium alone v/s Epithelium in co-culture	Macrophage alone versus Macrophage Coculture
MAP3772c	-1.266520914	0.635793715	0.121753283	0.176071006
MAP3773c	2.04894493	2.904894122	0.038697543	-0.159801049
MAP3774c	0.320455118	0.553603967	-0.337678336	0.02475353
MAP3776c	1.837440825	1.381992589	-1.015750241	0.336156445

Figure 9. **Differentially expressed genes identified by RNA-seq based on total read counts.** Fold Change ($log_{10}FC$) of gene expression in the LSP15 after MAP infection compared to the corresponding uninfected control, genes shown have a $P < 0.05$. Each bar color represents different genes on the genomic island: *MAP3772c* (blue); *MAP3773c* (orange); *MAP3774c* (black); *MAP3776c* (green). There is an up regulation of *MAP3773c* in the first two scenarios, suggesting a possible role of *MAP3773c* in MAP survival inside the epithelium.

3.5 DISCUSSION

The importance of a genomic island in the evolution and diversification of bacterial genomes has been studied in different organisms such as, *E. coli, S. enterica, H. pylori, M. genitalium, B subtillus,* and others (Orchman, Lawrence and Groisman, 2000). If a horizontal gene transfer (HGT) between unrelated species occurred and the new genomic region is retained by the recipient, it must confer an advantage and increase the fitness for colonizing hosts (Becq, Churlaud and Deschavanne, 2010). In MAP, previous characterization of LSP15 confirmed that they were acquired via HGT and its potential role in iron regulation and infection (Alexander, Turenne and Behr, 2009; Janagama *et al.,* 2010; Wang *et al.,* 2014). LSP15 contains 6 genes, one of them is *MAP3773c*, a putative ferric uptake regulator (Fur)-like transcription regulator confirmed by *in silico* amino acid similarity analysis. Confirming the transcriptional knockdown of expression of *MAP3773c* in the *MAP3776c*::Tn mutant (Fig. 4) provided us an alternate approach for initial investigations of the role of Fur in MAP that can likely lead to a better understanding of iron metabolism and its possible contribution on the MAP virulence.

In this present study, the disrupted strain was complemented with a copy of *MAP3776c* cloned under the hsp60 promoter. Although we were able to restore *MAP3776c* gene expression through complementation assay (Fig. 5), further investigation using *in vivo* and *ex vivo* models is necessary to ensure that the observed mutant phenotype is actually due to the loss of *MAP3776c* and not secondary mutations or polar effects that can be generated by the transposon (DeNicola *et al.,* 2015). In addition, it is been suggested that hsp60 promoters offers a strong expression that results in increased levels of the targeted gene (Sakthi and Narayanan, 2013; Seniya *et al.,* 2020). Hence, further analysis with the complemented strain or using an in-frame deletion of *MAP3773c* and its complemented counterpart would be necessary to conclusively establish the

phenotype and therefore the function of *MAP3773c*. Additional experiments by complementing one (*MAP3776c* alone) or all genes (*MAP3776c-MAP3772c*) in the putative operon will also need to be performed to reconfirm these findings.

In vitro models to study MAP have be shown to be useful to obtain insights into the host-pathogen interactions (Patel *et al.,* 2006). Here, using bovine mammary epithelia cells (MAC-T cells) we were able to demonstrate that lack of LSP15 impaired the ability of MAP to invade and survive inside the cells. While both cattle strains (MAP K10, MAP 1018) and the sheep strain (MAP 7565) invaded and persisted in high numbers throughout the 48 h experiment, *MAP3776c*::Tn appeared to be attenuated (Fig. 8). Wang and others (2014), using an *in vivo* mouse model, demonstrated that disruption of LSP15 lead to reduced bacterial fitness and were severely depleted after infection – corroborating our findings of *ex vivo* attenuation. Whether this phenotype was directly due to loss in function of *MAP3776c* and its downstream genes or due to polar effects needs confirmation.

Using a co-culture system where MAC-T cells and Monocyte-derived macrophages (MDMs) were cultured alone or together in a transwell system, wild type MAP phenotypes were analyzed (Fig. 9). While establishing this system, we have shown that MAP's first interaction within the host at the intestinal epithelium interface is a dynamic process that can be harnessed by the pathogen to achieve survival and dissemination within the macrophage (Lamont *et al.,* 2013). Our data showed that there is an upregulation of *MAP3773c* in epithelial cells in two different conditions (alone or in coculture with MDMs) (Fig. 9). These results suggest that MAP upregulates *MAP3773c* in MAC-T cells in co-culture with macrophages possibly as a likely survival strategy. In addition, Fur is known to regulate expression of genes involved in survival under iron-limiting conditions, including iron acquisition systems. Thus, we propose that while

the host creates a hostile environment by limiting iron, MAP upregulates *MAP3773c* to control iron homeostasis and improve its chances of survival. This will need confirmation using an in-frame deletion mutant of *MAP3773c*.

3.6 CONCLUSIONS AND FUTURE DIRECTIONS

In this chapter, we investigated the MAP-specific genomic island, LSP15. Our co-culture *in vitro* experiment recreating an early infection of MAP demonstrated that *MAP3773c* has a potential contribution in MAP survival inside epithelial cells. However, a deeper investigation is needed to validate these findings and determine the role of Fur in MAP infection. While using the *MAP3776c*::Tn provided us new insights in the MAP-iron interaction, there are several gaps that need to be addressed. The transposon mutagenesis used in this chapter was generated using the MycoMarT7 transposon, a vector that carries the highly active C9 Himar1 transposase outside the inverted repeats, resulting in a more stable mutant. However, the use of a transposon mutant is still not ideal when characterizing the function of a specific gene (Rubin *et al.,* 1999; Sassetti *et al.,* 2001; Wang *et al.,* 2014). The objective of a transposon is to move from one site to the next, even though the vector is considered stable, there is still the risk of instability in the phenotype of the strain carrying the transposon-derived mutation (Paustian and Kurtz, 1994). In addition, transposons can carry transcriptional terminators, thus when its "lands" in a gene in an operon, eliminates expression of all adjacent genes, mechanism often referred as polar effect (Hutchison III *et al.,* 2019). Furthermore, when using insertional mutagenesis there is also the damage done to the genome by the process of transposition itself as transposons are mobilized from chromosomal integration site that can result in frameshift mutations, splicing alterations or promoter disruptions (DeNicola *et al.,* 2015).

A more stable approach that can be used in order to confirm findings generated with the *MAP3776c*::Tn, is the construction of *MAP3773c* mutant with in-frame deletion using allelic exchange. Through homologous recombination, the cellular DNA repair mechanism aligns a targeting vector with its corresponding region of homology and cause recombination into the chromosome, the mutation generates a "loss of function" direct to a specific locus (Hall *et al.,* 2009). Homologous recombination takes advantage of a cell's own DNA repair machinery to replace the targeted locus with a replacement vector containing a drug selection marker flanked by two homology arms (Hall *et al.,* 2009).

In addition to homologous recombination, another system that can be used is the CRISPRi (clustered regularly interspaced short palindromic repeats interference) is a modification of the original CRISPR/Cas9 (crisper-associated protein 9) system, designed with a goal of gene silencing instead of knockouts. To date, there are two primary CRISPRi systems developed for *Mycobacterium*: codon-optimized dCas9 from *S. pyogenes* (Choudhary *et al.,* 2015, Singh et al., 2016) and dCas9 from *S. thermophilus* (Rock *et al.,* 2017). While both systems share the advantage of being anhydrotetracycline (ATc) inducible and fully reversible at any stage, the first using the more traditional *S. pyogenes* Cas9 shows less robust and specific gene knockdown than the second using *S. thermophilus* dCas9, with two vectors to select from to minimize toxicity from protein overexpression (Rathnaiah *et al.,* 2020).

The generation of mutations by molecular techniques is a useful way of advancing our understanding of iron regulation in MAP. A *MAP3773c* mutant will allow us to better understand Fur interaction with the network of genes that is deployed by MAP to maintain iron homeostasis. In addition, it would be interesting to investigate how Fur and IdeR work in a coordinate way to improve MAP survival inside the hostile environment or if Fur alone can improve MAP fitness.

REFERENCES

REFERENCES

Alexander, D.C., Turenne, C.Y., Behr, M.A. (2009) Insertion and deletion events that define the pathogen *Mycobacterium avium subsp. paratuberculosis.* J Bacteriol. 191(3):1018- 25. DOI: 10.1128/JB.01340-08

Bannantine, J.P., Hinnes II, M.E., Bermudez, L.E., Talaat, A.M., Sreevatsan, S., Stabel, J.R., Change, Y.F., Coussens, P.M., Barletta, R.G., Davis, W.C., Collins, D.M., Grohn, Y.T., Kapur, V. (2014a). A rational framework for evaluating the next generation of vaccines against *Mycobacterium avium* subspecies *paratuberculosis*. Frontiers in Cellular and Infection Microbiology. DOI:10.3389/fcimb.2014.00126

Becq, J., Churlaud, C., Deschavanne, P. (2010). A Benchmark of parametric methods for horizontal transfers detection. Plos one. 5(4). Doi:10.137/journalpone.0009989

DeNicola, G.M., Karreth, F.A., Adams, D.J., Wong, C.C. (2015). The utility of transposon mutagenesis for cancer studies in the era of genome editing. Genome Biology. 16:229. Doi: 10.1186/c13059-015-0794-y

Garcia, A.B., Shalloo, L. (2015) The economic impact and control of paratuberculosis in cattle. J. Dairy Sci. 98:5019–5039. doi:10.3168/jds.2014-9241

Guerinot, M. L. 1994. Microbial iron transport. Annu Rev Microbiol 48:743-72.

Hall, B., Limaye, A., Kulkarni, A.B. (2009). Overview: Generation of Gene knockout Mice. Curr Protoc Cell Biol. Doi:10.1002/0471143030.cn1912s44.

Hantke, K. (1981). Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol Gen Genet. 182, 288-92. DOI:10.1007/bf00269672

Hutchison III, C.A., Merryman, C., Sun, L., Assad-Garcia, N., Richter, R.A., Smith, H.O., Glass, J.I. (2019) Polar Effects of Transposon Insertion into a Minimal Bacterial Genome. Journal of Bacteriology. 185-19. doi:10.1128/JB.00185-19

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Janagama, H.K., Lamont, E.A., George, S., Bannantine, J.P., Xu, W.W., Tu, Z.J., Wells, S.J., Schefers, J., Sreevatsan, S. (2010) Primary transcriptomes of *Mycobacterium avium* subsp.

paratuberculosis reveal proprietary pathways in tissue and macrophages. BMC Genomics. 11, 561. doi: 10.1186/1471-2164-11-561.

Kugadas, A., Lamont, E.A., Bannantine, J.P., Shoyama, F.M., Brenner, E., Janagama, H.K., Sreevatsan, S. (2016). A *Mycobacterium avium* subsp. *paratuberculosis* predicted serine protease is associated with acid stress and intraphagosomal survival. Front Cell Infect Microbiol. 6, 85.

Lamont, E.A., Xu, W.W., Sreevatsan, S. (2013) Host-*Mycobacterium avium* subsp. *paratuberculosis* interactome reveals a novel iron assimilation mechanism linked to nitric oxide stress during early infection. BMC Genomics. 2164 DOI:10.1186/1471-2164-14-694

Lee, J., Helmann, J.D. (2007). Functional specialization within the Fur family of metalloregulators. Biometals. 20:485-499, doi:10.1007/s10534-006-9070-7

Li, L., Bannantine, J.P., Zhang, Q., Amonsin, A., May, B. J., Alt,D., Banerji, N., Kanjilal, S., Kapur, V. (2005) "The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*." Proc. Natl. Acad. Sci. 102:12344-12349. doi:10.1073/pnas.0505662102

Lombard, J.E. (2011) Epidemiology and Economics of Paratuberculosis Vet Clin Food Anim 27, 525-535. Doi:10.1016/j.cvfa.2011.07.012

MA, Z., Faulkner, M.J., Helmann, J.D. (2012). Origins of specificity and cross-talk in metal ion sensing by *Bacillus subtilis* Fur. Mol Microbiol. 86, 1144-55. DOI:10.1111/mmi.12049

Marino, R., Capoferri, R., Panelli, S., Minozzi, H., Strozzi, F., Trevisi, E., Snel, G.G.M., Ajmone-Marsan, P., William, J.L. (2017) Johne's disease in cattle: an in vitro model to study early response to infection of Mycobacterium avium subsp. paratuberculosis using RNA-seq. Mol. Immunology. 91, 259-271. doi:10.1016/j.imolimm.2017.08.017

Orchman, H., Lawrence, J.G., Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature (405)6764. doi:10.1038/35012500

Paustian, T. D., and R. S. Kurtz. 1994. Transposon mutagenesis of rhodobacter sphaeroides.

Pages 45-61, in Tested studies for laboratory teaching, Volume 15 (C. A. Goldman, Editor). Proceedings of the 15th

Workshop/Conference of the Association for Biology Laboratory Education (ABLE), 390 pages.

Park, K.T., Dahl, J.L., Bannantine, J.P., Barletta, R.G., Ahn, J., Allen, A.J., Jamilton, M.J., Davis, W.C. (2008) Demonstration of Allelic Exchange in the Slow-Growing Bacterium Mycobacterium avium subsp. paratuberculosis, and Generation of Mutants with Deletions at the pknG, relA, and lsr2 Loci. App and Env Micro. 1687-1695. Doi:10.1128/AEM.01208-07

Patel, D., Danelishvili, L., Yamazaki, Y., Alonso, M., Paustin, M.L., Bannantine, J.P., Meunier-Goddik, L., Bermudez, L.E. (2006). The Ability of *Mycobacterium*

avium subsp. *paratuberculosis* To Enter Bovine Epithelial Cells Is Influenced by Preexposure to a Hyperosmolar Environment and Intracellular Passage in Bovine Mammary Epithelial Cells. Infect. Immun. 74, 2849-2855. doi:10.1128/IAI.74.5.3849-2855.2006
Rathnaiah, G., Shoyama, F.M., Brenner, E.P., Zinniel, D.K., Bannantine, J.P., Sreevatsan, S., Chacon, O., Barletta, R.G. (2020). Molecular Genetics of Mycobacterium avium subsp. paratuberculosis. Paratuberculosis: Organism, Disease, Control. CABI

Sakthi, A., Narayanan, S. (2013). The *lpqS* knockout mutant of *Mycobacterium tuberculosis* is attenuated in Macrophages. Microbiological Research. 168:404-414. Doi:10.1016/j.micres.2013.02.007.

Seniya, S.P., Yadav, P., Jain, V. (2020) Construction of E. coli—Mycobacterium shuttle vectors with a variety of expression systems and polypeptide tags for gene expression in mycobacteria. Plos One. DOI:10.1371/journal.pone.0230282

Stratmann, J., Strommenger, B., Goethe, R., Dohmann, K., Gerlach, G.F., Stevenson, K., Li, L., Zhang, Q., Kapur, V., Bull, T. (2004) A 38-kilobase pathogenicity island specific for *Mycobacterium avium* subsp. *paratuberculosis* encodes cell surface proteins expressed in the host. Infect. Immun. 72, 1265–1274. DOI: 10.1128/IAI.72.3.1265-1274.2004

Wang, J., Moolji, J., Dufort, A., Staffa, A., Domenech, P., Reed, M.B., Behr, M.A. (2016) Iron acquisition in *Mycobacterium avium* subsp. *paratuberculosis.* J. Bacteriol. 198:857–866. DOI:10.1128/JB.00922-15

Wang, J., Pritchard, J.R., Kreitmann, L., Montpetit, A., Behr, M.A. (2014) Disruption of *Mycobacterium avium* subsp. *paratuberculosis*-specific gene impairs in vivo fitness. BMC Genomics. 15:415. DOI:10.1186/1471-2164-15-415

Whittington, R.J., Marsh, I., Turner, M.J., McAllister, S., Choy, E., Eamens, G.J., Marshall, D.J., Ottaway, S. (1998). Rapid detection of *Mycobacterium paratuberculosis* in clinical samples from ruminants and in spiked environmental samples by modified BACTEC 12B radiometric culture and direct confirmation by IS900 PCR. J Clin Microbiol. 36, 701-7.

Zhu, X., Tu, Z.J., Coussens, P.M., Kapur, V., Janagama, H., Naser, S., Sreevatsan S. (2008) Transcriptional analysis of diverse strains *Mycobacterium avium* subspecies *paratuberculosis* in primary bovine monocyte derived macrophages. Microbes and Infection. 1274-1282. doi:10.1016/j.micinf.2008.07.025

CHAPTER FOUR: ELUCIDATING THE REGULON OF A FUR-LIKE PROTEIN IN *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS* (MAP)

4.1 ABSTRACT

Intracellular iron concentration is tightly regulated to maintain cell viability. Iron plays important roles in electron transport, nucleic acid synthesis and oxidative stress. A *Mycobacterium avium* subsp. *paratuberculosis (MAP)*- specific genomic island carries a putative metal transport operon that includes *MAP3773c*, which encodes a Fur-like protein. Although well characterized as a global regulator of iron homeostasis in multiple bacteria, the function of Fur (ferric uptake regulator) in MAP is unknown as this organism also carries, IdeR (iron dependent regulator), a native iron regulatory protein specific to mycobacteria. Computational analysis using PRODORIC identified 23 different pathways involved in respiration, metabolism and virulence that were likely regulated by *MAP3773c*. Thus, chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) was performed to confirm the putative regulon of *MAP3773c* (Fur-like protein) in MAP. ChIP-Seq revealed enriched binding to 58 regions by Fur under iron replete and deplete conditions, located mostly within open reading frames (ORF). Three ChIP peaks were identified in genes that are directly related to iron regulation: *MAP3638c* (hemophore-like protein), *MAP3736c* (Fur box) and *MAP3776c* (ABC transporter*).* A Fur box consensus sequence was identified, and binding specificity and dependence on Mn^{2+} availability was confirmed by a chemiluminescent electrophoresis mobility shift assay (EMSA). The results confirmed that *MAP3773c* is a Fur ortholog that recognizes a 19-bp DNA sequence motif (Fur box) and it is involved in metal homeostasis. This work provides a regulatory network of MAP Fur binding sites during iron replete and deplete conditions and highlights unique properties of Fur regulon in MAP.

4.2 INTRODUCTION

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of Johne's disease (JD) in ruminants, a chronic and incurable chronic enteritis characterized by persistent diarrhea that leads to malnutrition and muscular wasting (Rathnaiah *et al.,* 2017). JD is present worldwide and imposes significant economic losses to the dairy industry. (Garcia and Shalloo, 2015). Unfortunately to date, reliable JD diagnostics are still lacking. Culture of *MAP* from feces has been the most reliable method for diagnosis of JD, however MAP requires eight to sixteen weeks to produce colonies in culture, presenting a major hurdle to diagnosis (Bannantine *et al.,* 2002). Unlike other mycobacteria, MAP has special iron requirements. For optimal growth *in vitro*, it requires supplementation of the siderophore mycobactin J. Whole genome sequencing of MAP K10 provided a potential explanation for this dependency, revealing a truncation of the *mbtA* gene, with MAP making a protein that is 151-156 amino acid shorter than *M. tuberculosis* or *M. avium* (Li *et al.*, 2005). It has been suggested that this truncation impairs the production of mycobactin from the *mbtA-J* operon (Li *et al.*, 2005; Wang *et al.,* 2014). Despite this truncation, Zhu *et al.* (2008) showed that MAP is still able to transcribe mycobactin synthesis genes inside macrophages. To corroborate these findings, Janagama *et al*., described the upregulation of several genes responsible for iron acquisition in infected tissues, including genes responsible for mycobactin biosynthesis (Janagama *et al.*, 2010).

Iron is vital to fundamental biological processes; however, high intracellular concentrations of free iron are toxic to bacteria. As such, cells have developed tightly regulated processes for intracellular metal homeostasis (Eckelt *et al.*, 2014). Bacteria control metal homeostasis by activating a set of genes regulated by metal-sensing transcription factors known as metalloregulatory proteins (Chandrangsu *et al.*, 2017). In prokaryotes, there are two major

families of metalloregulators: diphtheria toxin (DtxR), and ferric uptake regulator (Fur) (Hantke, 2001). In 2009, Janagama and others identified and characterized *MAP2827*, an iron-dependent regulator (IdeR) in MAP. A member of the DtxR protein family, IdeR is involved in regulatory mechanisms to acquire, store or prevent excess accumulation of iron. The authors were able to confirm that *MAP2827* was in fact IdeR and regulates genes involved in iron acquisition (*mbtB*) and iron storage (*bfrA*) (Janagama *et al.,* 2009). However, *in vitro* iron stress showed that IdeR regulation is strain dependent – while IdeR from MAP cattle strain K10 regulates mycobactin synthesis and storage genes similar to IdeR from *M. tb*, IdeR from MAP sheep strain S397 shows deficiency in iron storage function, resulting in a strain more sensitive to iron fluctuations (Janagama *et al.,* 2010).

In addition to IdeR, MAP genome contains putative metal transport MAP-specific operon, large genomic polymorphisms (LSPs) 15, that includes a Fur-like transcriptional regulator, *MAP3773c* (Alexander *et al*., 2009). First identified in *E.coli,* Fur has been shown to respond to iron replete conditions to repress gene expression and allow sufficient concentration of intracellular iron for essential iron related activities (Hantke, 1981; Bagg and Neilands, 1987; Lee and Helmann, 2007). Similar to several representative of Fur family member, Fur protein requires binding of a divalent metal ion, either Fe^{2+} or Mn^{2+}, for DNA-binding activation (Mills and Marletta, 2005; Lee and Helmann, 2007; Chandrangsu *et al.*, 2017). Fur protein generally binds to a 19-bp inverted repeat sequence known as a "Fur box" (GATAATGATwATCATTATC; w = A or T), within the promoter of the regulated genes (Escola, Pérez-Martin and Lorenzo, 1999). In MAP, functional genomics suggested three Fur boxes located in a 38-kb *MAP*-specific genomic island (LSP14) (Stratmann *et al.,* 2004; Alexander *et al*., 2009). *MAP* genome includes a total of six specific genomic insertions: LSP4, LSP11, LSP12, LSP14, LSP15 and LSP16 (Alexander *et.*

al.,2009). As these islands are not presented in any other mycobacteria, it has been proposed and confirmed that they were acquired via horizontal gene transfer (Alexander *et. al*.,2009; Wang *et al*., 2016). Furthermore, LSP14 and LSP15 encodes several predicted genes involved in metal uptake systems.

To date, different studies have characterized Fur family members in MAP: FurA and FurB, also known as Per (peroxidase stress response) and Zur (zinc uptake repressor) respectively, however no information about the potential role of *MAP3773c,* Fur-like element, has been described *(*Eckelt *et al.*, 2014; Eckelt *et al.*, 2015).

As a key virulence determinant, iron regulation in MAP and its role in pathogen survival and infection are important areas of research that may lead to advances in ability to improve culturing methods. To further elucidate the mechanisms of iron homeostasis in MAP*,* we investigated the putative function of the Fur-like gene (*MAP3773c*) in iron homeostasis *in vitro*. We applied *in vivo* ChIP-seq to confirm binding of MAP Fur as a transcription factor and to identify the regulon of genes under its control.

4.3 MATERIAL AND METHODS

4.3.1 BACTERIAL STRAINS

MAP K10 strain was grown at 37°C without shaking in Middlebrook 7H9 supplemented with 10% OADC (oleic acid, dextrose, catalase) enrichment medium (Thermo Scientific, Waltham, MA), 0.05% tween 80 and 2 mg of ferric mycobactin J (Allied Monitor Inc, Fayette, MO) per liter. Antibiotics (μg/ml: kanamycin, 20; hygromycin, 100; streptomycin, 20) were added when necessary. Competent *E. coli* BL21(DE3) (EMD Biosciences, Madison WI) and *E. coli* TOP10F cells (Invitrogen, Carlsbad, CA) were grown in LB medium 37°C with shaking at 200 RPM.

4.3.2 PROTEIN EXPRESSION

To express MAP Fur protein, competent *E. coli* BL21(DE3) (EMD Biosciences, Madison WI) carrying *MAP3773c* on pET-24b(+) were growing in LB medium with 30µg/mL kanamycin. Cultures were kept at 37°C with shaking at 200 RPM for 4 h aerobic growth, until OD_{600} of 0.4 was obtained. Then protein expression was induced with addition of 0.1 M IPTG and shaking at 37°C for an additional 2 h. The expressed *MAP3773c* was extracted using B-PER (Bacterial Protein Extraction Reagent; Pierce Biotechnology, Rockford, IL), followed by purification using HisPur Ni-NTA resin columns per manufacturer's protocol (Pierce Biotechnology, Rockford, IL). Purified protein was analyzed by SDS-PAGE and Western Blot using standard methods described previously (Bannantine and Paustin, 2006). The target band identified from the SDS-PAGE was excised for LC-MS/MS at Michigan State University Proteomics Facilities. Raw data were analyzed using Scaffold (Proteome Software, Portland, OR).

4.3.3 WESTERN BLOTTING

MAP K10 were cultured until OD_{600} of ~0.5 was reached. For iron starvation, cultures were treated with 2,2'-bipyridyl (DIP, 200 µM final) for 2 hours shacking at 200 rpm at 37 °C. Cells from iron replete and deplete conditions were washed with 1x PBS and resuspended in freshly made buffer lysis buffer (20 mM HEPES; 50 mM KCl; 0.5 mM DTT; 10% glycerol; mini protease inhibitor), followed by cell lysing with MagNA Lyser (Roche Diagnostic, Sandhofer, Germany). For enrichment of Fur protein, samples were subjected to immunoprecipitation. Samples were incubated overnight with antibody for Fur detection at 4 °C on a rotating platform followed by 2 hours incubation (0.5 hour at 4 °C and 1.5 hours at RT) on a rotating platform. Samples were washed 2 times with IPP150 buffer (10 mM Tris-HCl; 150 mM NaCl, 0.1%

NP40) and 2 times with 1x TE (0.05 M Tris-HCl; 10mM EDTA) buffer. Beads were resuspended in elution buffer and incubated at 65 °C for 15 min. The samples were subjected to SDS-PAGE and transferred to Nitrocellulose Membrane, 0.2 μm (Bio-rad Laboratories, Hercules, CA). Custom-made antibody that binds the MAP Fur protein (Genscript, Piscataway, NJ) was used as primary antibody. Anti-Rabbit IgG (whole molecule)–Peroxidase antibody produced in goat (Sigma-Aldrich, St Louis, MO) was used as secondary antibody. The membrane was visualized with ChemiDoc MP Imaging System (Bio-rad Laboratories, Hercules, CA).

4.3.4 COMPUTATIONAL PREDICTION OF FUR-REGULATED GENES

Virtual Footprint, part of The Prokaryotic Database of Gene Regulation (PRODORIC) (Münch *et al*. 2005), was used for prediction of Fur binding site. *MAP* K-10 genome was used as input DNA sequence, Fur box motif from *E. coli* was used as Position Weight Matrix and searches were limited to -300 to + 100 bases of each predicted ORF.

4.3.5 CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY SEQUENCING (CHIP-SEQ)

ChIP-enriched DNA samples were performed following the protocol developed by Jaini *et al.* (2013) using a custom-made antibody that binds the MAP Fur protein (Genscript, Piscataway, NJ). MAP K10 culture with OD_{600} of ~0.5 was used to generate ChIP-DNA. In order to avoid false positive, input DNA was used as control, this sample did not have ChIP enrichment. For iron starvation, cultures were treated with 2,2'-bipyridyl (DIP, 200 μM final) for 2 hours shacking at 200 rpm at 37 °C. Cells from iron replete and deplete conditions were washed with

1x PBS. Formaldehyde was added at a final concentration of 1% and incubated at RT for 20 min in a platform rocker. Crosslinking was quenched by adding 250 mM of glycine and incubating for 15 min. Cells were washed two times with ice-cold 1x PBS and resuspended in freshly made lysis buffer (20 mM HEPES; 50 mM KCl; 0.5 mM DTT; 10% glycerol; mini protease inhibitor), followed by cell lysing with MagNA Lyser (Roche Diagnostic, Sandhofer, Germany). Cells suspension were sonicated using Covaris M220 Focused-ultrasonicator (Covaris, Inc., Woburn, MA) for 18 min; 75.0 peak power; 20.0 duty factor and 200 cycles/ burst. Samples were incubated overnight with antibody for Fur detection at 4 °C on a rotating platform followed by 2 hours incubation (0.5 hour at 4 °C and 1.5 hours at RT) on a rotating platform. Samples were washed 2 times with IPP150 buffer (10 mM Tris-HCl; 150 mM NaCl, 0.1% NP40) and 2 times with 1x TE (0.05 M Tris-HCl; 10mM EDTA) buffer. Beads were resuspended in elution buffer and incubated at 65 °C for 15 min. 1 mg/ml of Proteinase K was added to each sample and incubated at 37 °C for 2 hours and transferred for 65 °C for overnight incubation. The DNA purification was performed using AmPure^{xp} beads per manufacturer's protocol (Beckman Coulter, Indianapolis, IN). Samples quality were analyzed by an Agilent 2100 Bioanalyzer (Agilent technologies, Santa Clara, CA).

4.3.6 CHIP-SEQ LIBRARY CONSTRUCTION AND SEQUENCING

DNA fragments ~300bp were selected for library preparation and sequencing libraries prepared using NEXTflex^{tm} ChIP-seq kit (PerkinElmer, Austin, TX) as per manufacturer's protocol. Prior and post library construction, chromatin immunoprecipitation products were quantified using Qubit fluorometer (Invitrogen, Carlsbad, CA) and Agilent 2100 Bioanalyzer (Agilent technologies, Santa Clara, CA). ChIP DNA replicates were pooled and sequenced.

Approximately 20M reads per sample were generated, providing 150-1000 depth of coverage. Sequencing was performed by ACGT, Inc (Chicago, IL).

4.3.7 CHIP-SEQ DATA ANALYSIS

All analysis was done using CLC Genomics Workbench software 12.0 (QIAGEN, Aarhus, Denmark). Raw data generated from ChIP-seq were trimmed and mapped to the reference MAP K10 genome (NCBI accession number NC_002944). Using CLC shape-based peak caller, ChIP-enriched DNA were aligned onto Input DNA (no ChIP enrichment), when the sequence coverage of a genomic region in the enriched DNA exceeded the Input DNA, a ChIP peak score was called. A list of all ChIP peaks with their respective P-value was generated. The threshold for signal to noise ratio (ChIP-enriched DNA versus no enriched) was set based on false-discovery rate (FDR) value equal or smaller than 10^{-50}. FDR was calculated using Bonferroni correction on R software based on the P-value generated by CLC Workbench.

4.3.8 MOTIF DETECTION

A Fur binding motif was generated using Find Individual Motif Occurrences (FIMO), part of the MEME suit (Grant, Bailey & Noble, 2011), for all *in-vivo* binding sites identified in ChIP-seq analyses. A p-value of ≤ 0.001 was defined as statistical threshold for Fur binding motifs.

4.3.9 ELECTROPHORETIC MOBILITY SHIFT ASSAY (EMSA)

Physical binding of *MAP3773c* to the promoter sequences of *MAP* Fur box 1 (*MAP3736*c) was carried out by EMSA. Promoter sequences containing Fur box motifs were amplified using 5' biotin-labeled primer via PCR. Purification of PCR products were done using the QIAquick PCR

Purification kit (Qiagen, Germantown, MD). Recombinant *MAP* Fur protein was expressed as stated above. Binding reaction: 1x Binding Buffer (50 mM TrisHCl; 25% Glycerol; 10 μg/ml Salmon testes DNA; 250 nM NaCl; 5mM DTT; 250 μg/ml BSA; nuclease free water), 10 mM $MnCl_2$, 0-10 nM *MAP* Fur protein, 0-4 pmol Unlabeled DNA, 20 fmol labeled DNA. The reactions were incubated for 30 min at room temperature followed by electrophoresis in a 5% native polyacrylamide gel (40% 19:1 Acrylamide; 50% Tris-Acetate (TA) buffer; 50% glycerol; 10% Ammonium persulfate (APS); 6% TEMED) using 1x TA Buffer (1 M Tris acetate, 0.5 M Glacial acetic acid) as running buffer. After electrophoresis, gels were transferred onto a Biodyne B Nylon membrane (Pierce, Biotechnology, Rockford, IL) and reactions were detected using chemiluminescence-based nucleic acid detection kit (Pierce, Biotechnology, Rockford, IL).

4.4 RESULTS

4.4.1 GENOME-WIDE ANALYSIS OF FUR REGULON

Using computational prediction, PRODORIC (Münch *et al*. 2005), 26 different pathways involved in respiration, metabolism and virulence were identified as likely regulated by *MAP3773c* (Fig. 10).

To confirm the findings from the *in silico* analysis and determine which genes are regulated by Fur in MAP, chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) was performed. A custom-synthesized anti-Fur antibody capable of detecting the MAP Fur protein in its native form in MAP K10 (Fig. 13) was used to generate ChIP binding profiles for MAP K10 cultured under iron-replete and deplete conditions (Fig. 14). ChIP peaks were called when the sequence coverage of genomic regions in the different treatments is enriched when compared to ChIP-seq control sample where the immunoprecipitation step was omitted (Strino & Lappe

2016). Input DNA (no ChIP enrichment), had 34,907,295 (79.02% coverage against the MAP K10 genome) uniquely-mapped reads while ChIP-enriched DNA from iron-replete and iron-deplete conditions had 22,566,602 (55.66%) and 4,299,792 (14.73%) mapped reads respectively.

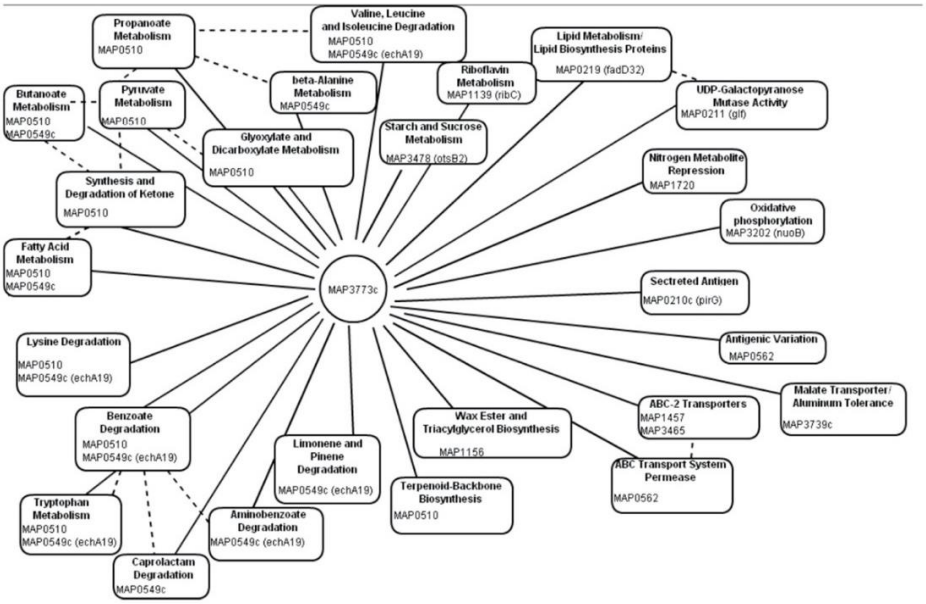


Figure 10. ***In silico* analysis of Fur regulon**. Using PRODORIC for MAP K10 genome analysis to detect putative Fur binding and predict pathways regulated by *MAP3773c*. Solid lines represent pathways directly regulated by *MAP3773c*. Dashed lines indicated inter-related pathways.

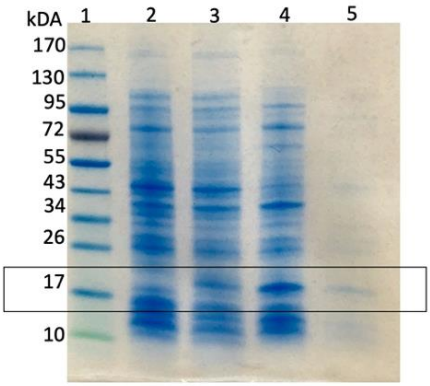


Figure 11. **Identification of MAP Fur protein.** Coomassie stain of SDS-PAGE analysis of Fur expression in E. coli system. Lane 1. Protein ladder; Lane 2. BL21(DE3) carrying empty vector pET24b(+); Lane 3. BL21(DE3) carrying MAP3773c on pET24b(+); Lane 4. BL21(DE3) carrying MAP3773c on pET24b(+) with addition of IPTG; Lane 5. Purified recombinant MAP3773c

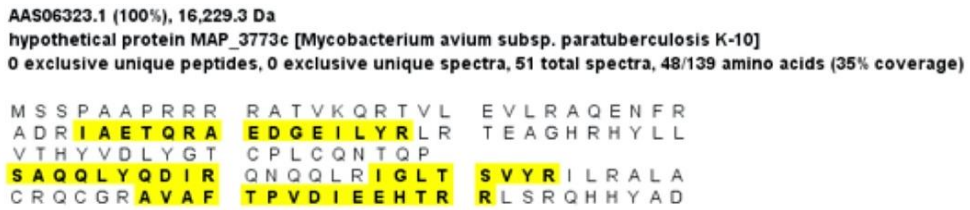


Figure 12. **Scaffold analysis of LC-MS/MS.** Data from excised band from lane 5 showing peptides hits (yellow highlights) to 35% of complete MAP Fur sequence.

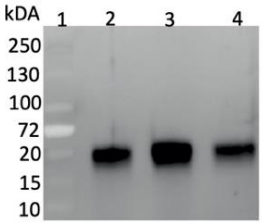


Figure 13. **Western blot of Fur protein.** Data showing immunoprecipitation of Fur protein by anti-Fur antibody from MAP K-10 cultured under iron-replete and deplete condition. Lane 1. Protein ladder; Lane 2. Pull down of MAP K-10 cultured under iron-replete condition using 1 μg of anti-Fur antibody; Lane 3. Pull down of MAP K-10 cultured under iron-replete condition using 2 μg of anti-Fur antibody; Lane 4. Pull down of MAP K-10 cultured under iron-deplete condition using 1 μg of anti-Fur antibody.

Applying a *P*-value at ≤ 0.001, the ChIP-seq assay identified nine Fur-binding sites out of 14 previously predicted by PRODORIC. ChIP-seq analysis revealed a total of 5,381 and 4,960 binding sites of Fur protein in the MAP K10 genome (signal to noise ratio) under iron-replete and iron-deplete conditions, respectively (Fig. 14).

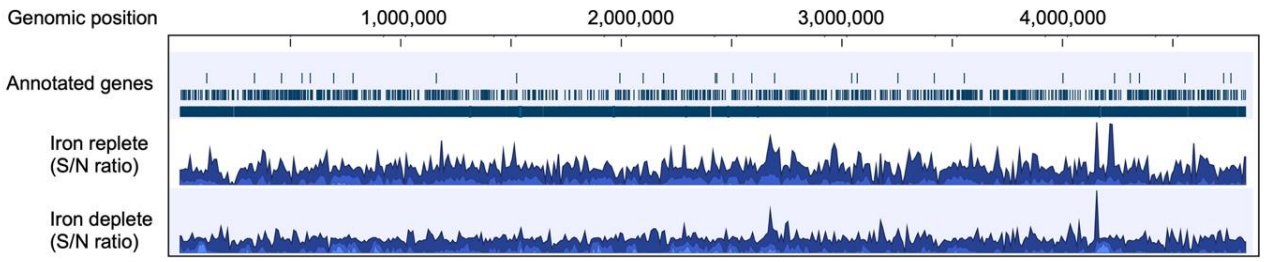


Figure 14. **Overview of the mapped sequences within the reference genome MAP K-10 under iron replete and deplete conditions generated by CLC genomic workbench.** After mapped onto the reference genome, iron replete and deplete samples were compared to control (input DNA) and signal-to-noise (S/N) ratio for peak calling was generated. Fur specifically binds various genomic loci under both conditions, but most of the ChIP peaks showed higher binding sites under iron-replete condition. Arrows indicate regions where ChIP peaks are associated with iron regulation.

Applying a false discovery rate (FDR) at ≤ 10^{-50}, under iron replete conditions, a total of 43 significantly enriched regions were identified on the K10 genome (Table 1). Peaks were either localized between open reading frames (ORF) (27%; intergenic regions) and within annotated genes (73%). In contrast under chelation treatment (iron-depletion), 11 enriched regions were identified (Table 2), all showing binding sites within ORFs. Four ChIP peaks were present under both iron replete and deplete conditions simultaneously (Table 3). Diverse functions are encoded by genes where Fur bound on the MAP K10 genome: cell wall synthesis, energy metabolism, respiration and transcriptional/ translation regulation. Out of 58 genes (FDR ≤ 10^{-50}) from both conditions (Table 1, 2 and 3), 11 are annotated as hypothetical proteins, two described as pseudogenes (Table 4) and three ChIP peaks are associated with iron regulation: *MAP3638c, MAP3736c and MAP3776c.*

Interestingly, Fur bound upstream of *MAP3776c*, an ABC transporter, only under iron-replete condition and binding to *MAP3638c* (hemophore-like protein) was identified only under iron-deplete conditions (Fig. 15 and 16).

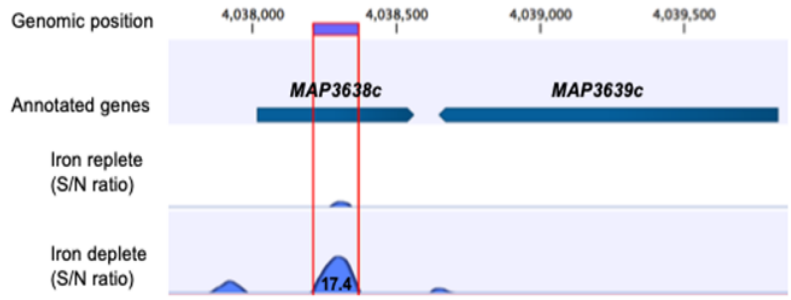


Figure 15. **Applying FDR ≤ 10^{-50}, there are three ChIP peaks associated with iron regulation**. MAP Fur protein binds to the region of MAP3638c, however only under iron-deplete condition binding is statically significant with peak score of 17.4 (MAP3638c).

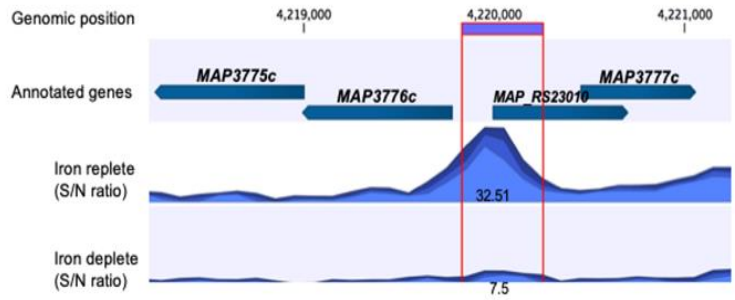


Figure 16. **MAP Fur binding to the region of MAP3776c under iron replete condition**. Under iron-replete condition there is a strong binding of MAP Fur to the region of MAP3776c represented by a peak score of 32.51.

4.4.2 FUR BINDS TO FUR BOX MOTIF UNDER IRON REPLETE OR DEPLETE CONDITION

Fur box consensus sequence was identified in ChIP-seq data using MEME-ChIP (Fig. 17). FIMO (Find Individual Motif Occurrences) analysis identified 15 occurrences of Fur box motif (*P*-value ≤ 0.001), 12 of them presented under iron replete conditions and three under iron deplete condition (Table 5).

Figure 17. **MAP Fur box analysis**. The most significant motif derived from ChIP-seq binding sequence using MEME. Height of each letter represents the relative frequency of each base at a different position in the consensus sequence.

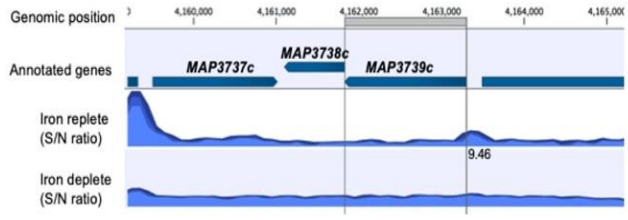


Figure 18. **A zoom-in of the MAP Fur box 3.** Under both iron conditions there is no binding of MAP Fur to the region of Fur box 3 (*MAP3739c*). ChIP peak (9.46) outside the ORF has FDR higher than the threshold of FDR ≤ 10^{-50}.

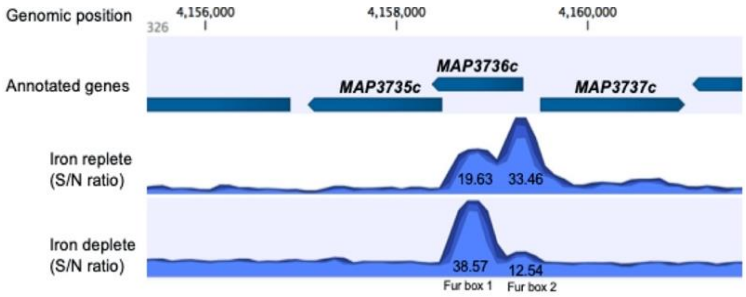


Figure 19. **The enriched region of MAP Fur binding onto Fur Box 1 and 2 identified by ChIP-seq.** ChIP peak showed higher occupancy under iron-deplete condition in the Fur Box 1 region. S/N denotes the signal-to-noise ratio for peak calling generated by CLC software.

From previous studies, it is known that the MAP K10 genome contains three Fur box motifs (Stratmann *et al.,* 2004). However, data from ChIP-seq showed that the Fur protein does not show significant binding (FDR ≥ 10^{-50}) to the region of Fur box 3 (*MAP3739c*) (Fig. 18). The highest peak score from all ChIP-seq data was observed within and just upstream of *MAP3736c*, located on LSP14, *MAP*-specific genomic island (Alexander *et al.,* 2009). Within *MAP3736c* (located between nucleotides 4158368 and 4159327), there are two putative Fur Boxes: Fur box 1 (located between nt 4158681 and 4158966 of the genome), and Fur box 2 (located between nt 4159132 and 4159456) (Stratmann *et al.,* 2004). ChIP-seq analysis showed high binding in both regions, confirming the exact location (Fig. 19). When intracellular Fe^{2+} was depleted by the addition of 2,2-dipyridyl, MAP Fur bound with higher affinity to Fur box 1 region (peak score = 38.57) in contrast to a lower binding score for Fur box 2 (peak score = 12.54). While under replete conditions, where MAP was grown in a complete media, the opposite was observed, a lower MAP Fur binding in the Fur box 1 region (peak score = 19.63) and a higher peak in Fur box 2 region (peak score = 33.46).

4.4.3 VALIDATION OF MAP FUR-BINDING

To confirm binding to the Fur promoter region, biotinylated or unlabeled PCR fragment including Fur box 1 identified by ChIP-seq was amplified and subjected to an electrophoretic mobility shift assay (EMSA) using purified MAP Fur protein (Fig. 11 and 12).

Titration of Fur protein in the presence of Mn^{2+} and 20 fmol of DNA showed that binding is dose-dependent, as the Fur concentration was increased there was an increase of binding activity (Fig. 20). However, in the absence of Mn^{2+}, Fur binding to DNA was not as efficient as in the presence of Mn^{2+} (Fig. 21). Furthermore, DNA-protein complex was specific to Fur binding site, as showed in the competition assay (Fig. 22), adding different concentration of excess unlabeled Fur box 1 probe competed with and abrogated labeled Fur box 1 probe binding to Fur protein.

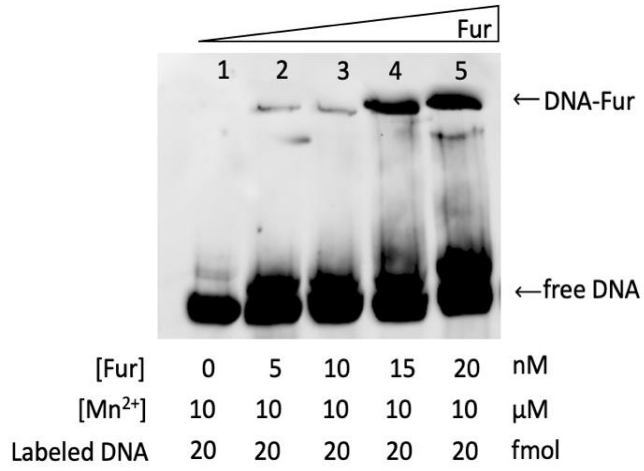
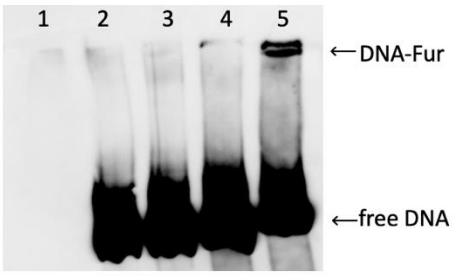
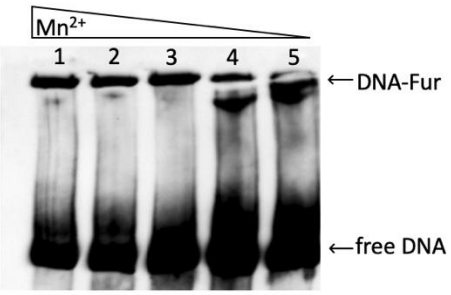


Figure 20. **EMSA analysis of MAP Fur binding to Fur box consensus DNA.** Binding activity is represented by band intensity. 20 fmol of MAP DNA including the Fur Box 1 consensus biotin labeled was run in a 5% native polyacrylamide gel with different concentration of MAP Fur protein and Mn^{2+}; Protein-DNA binding is dose-dependent: titration of purified MAP Fur protein shows an increase of binding activity as more protein is added to the system.



Labeled DNA	0	5	10	20	20	fmol
Unlabeled DNA	4	3	2	1	0	pmol
[Fur]	20	20	20	20	20	nM
[Mn^{2+}]	10	10	10	10	10	µM

Figure 21. **EMSA analysis of Fur – Fur box binding activity is more efficient in the presence of Mn^{2+}.** No addition of Mn^{2+} (Lane 5) binding occurs with a lower band intensity when compared to sample with Mn^{2+} (Lane 1-4).



[Mn^{2+}]	20	15	10	5	0	µM
[Fur]	20	20	20	20	20	nM
Labeled DNA	20	20	20	20	20	fmol

Figure 22. **MAP Fur binding is specific to Fur box consensus.** Competitive EMSA, Fur protein was incubated with either biotin-labeled DNA probe or unlabeled DNA probe or with both. Biotin-labeled probe was detected using chemiluminescent-based nucleic acid detection kit. Addition of unlabeled DNA affects binding activity, showing binding specificity.

4.5 DISCUSSION

In this study, a full characterization of the Ferric Uptake Regulator in MAP was performed. Fur and its involvement in iron homeostasis is well known in bacteria such as *E. coli*, *B. subtilis* and *S. Typhimurium.* This protein has been shown to work as a repressor, by blocking RNA polymerase binding to the promoter region of genes involved in iron homeostasis by repressing transcription (Escolar *et al.,* 1997); but can also work as an activator by positively regulating

gene expression in response to iron through indirect mechanism involving repression of small regulatory RNA (Delany *et al.,* 2001; Masse & Gottesman 2005). The current study confirmed by Western Blot (Fig. 11) and mass spectrometry (Fig. 12) that *MAP3773c* encodes a Fur-like protein in MAP. A regulatory network of MAP Fur binding sites was identified using three independent approaches: 1. *In-silico* (PRODORIC); 2. *In vivo* (ChIP-seq) and 3. *In vitro* (EMSA). *In vivo* and *in vitro* analyses established that Fur binding was responsive to iron availability.

ChIP-seq analysis expanded the number of MAP Fur binding sites, from 14 genes predicted by PRODORIC to 58 enriched binding regions (FDR $\leq 10^{-50}$). Binding locations were distributed almost evenly between intragenic and intergenic regions. While binding of Fur in intragenic regions refute the definition of a transcriptional factor (Browning and Busby, 2004), recent ChIP-seq studies with *M. tuberculosis, E.coli, Salmonella* and *Corynebacterium* reported intragenic TF binding that play critical roles in transcription and significantly affect regulation of gene expression (Knapp *et al.,* 2015; Fitzgerald *et al.,* 2014; Dillon *et al.,* 2012). Additionally, during characterization of the Fur regulon in *Pseudomonas syringae*, Butcher and others (2011) did not observe general differences between Fur binding to intergenic and intragenic sites. Both showed comparable binding affinity in *P. syringae* suggesting that, although 100% of MAP Fur binding under iron-deplete conditions are located in intragenic regions, MAP Fur can be biologically active and able to bind to specific DNA sequences to control gene expression.

Iron regulation by Fur in MAP appears to be more complex than the classic model, where Fur acts as a repressor when sensing high intracellular Fe^{2+}. It then forms the Fur-Fe^{2+} complex and binds to the Fur box sequence, which enables Fur transition from its inactive (*apo*-) to its activated (*holo-)* form (Hantke, 2001; Helmann, 2014). Additionally, data from the present study

showed that *MAP* uses Fur in the absence of intracellular Fe^{2+}, a process known as *apo*-regulation. In low iron conditions *apo*-Fur protein bind to the promoters of its target genes and regulates transcription (Miles *et al.,* 2010).

The complexity of Fur regulation can be exemplified in the ChIP peak of *MAP3736c*, where *apo*-Fur binds to Fur box 1 under iron-deplete conditions and *holo*-Fur binds to Fur box 1 and 2 under iron-replete condition. The physiological significance of *apo*-Fur binding in MAP is unclear, however previous studies with *H. pylori* showed that when iron levels are low, genes responsible for iron storage are repressed by *apo*-Fur (Bereswill *et al.*, 2000). Furthermore, additional studies in *C. jejuni* showed that gene expression controlled by Fur was decreased in the wild type strain under iron depleted condition and, in a Fur knockout strain, expression was increased (Holmes *et al.,* 2005), indicating that *apo*-Fur plays an important role in iron metabolism. Corroborating this result, ChIP-seq analysis identified *apo*-Fur binding to *MAP3638c*, only under iron starvation. MAP3638c is a hemophore-like protein, suggesting that *MAP* likely uses heme as an additional iron source as previously described in *M. tuberculosis* (Tullius *et al.,* 2011).

Finally, to confirm and validate Fur-Fur box1 binding, an EMSA using PCR amplification of ChIP-seq-identified Fur box 1 and purified Fur-like protein (*MAP3773c)* was performed. The binding was dependent on the availability of Mn^{2+}, a common surrogate metal which, unlike Fe^{2+}, is stable in the presence of oxygen but promotes DNA binding and adopts the same coordination geometry as Fe^{2+} (Butcher *et al.,* 2011). Additionally, a competitive gel shift assay confirmed specificity of *MAP* Fur binding to the Fur box 1 region. Taken together, the identification of consensus Fur box by ChIP-seq peaks combined with data from EMSA, confirms that iron regulation in MAP is also mediated by a Fur homolog that recognizes the 19bp DNA sequence, known as the Fur box.

In this current study, we were not able to confirm Fur box 3 (*MAP3739c*) region as a binding site for Fur protein as described by Stratmann and others (2004). Computational methods as used by the group predicted binding sites relying on data available 15 years ago, which was likely incomplete. Further, most computational predictions of transcription factors binding are prone to false discovery and need to be validated (Karimzadeh and Hoffman, 2018). By using directly and quantitatively sequencing in combination with specific antibody, as used in this currently study, ChIP-seq method provides a powerful strategy for identifying *in vivo* binding sites across the entire genome (Collas, 2019).

4.6 CONCLUDING REMARK AND FUTURE DIRECTIONS

In this work, we characterized *MAP3773c*, the ferric uptake regulator in MAP, using ChIP-seq. A genomic view of the MAP Fur regulatory network was identified, and several putative binding sites involved during iron replete and deplete conditions were discovered. Although this study is not a full description of the Fur regulon, our findings indicate that MAP Fur is a global regulator that recognizes many target sites in the genome, either by *apo*- or *holo*- Fur.

Further analysis of the complete MAP Fur regulon will combine ChIP-seq data analysis from this work with another genome-scale experiment will provide a full understanding of direct or indirect roles of Fur in response to iron availability. To have a complete understanding of the MAP iron stimulon model, future studies will involve the basic understanding of Fur-IdeR interactions and how one or the other may be functional in MAP under a variety of *in vivo* and environmental conditions.

Table 1. List of genes regulated by Fur under iron replete condition FDR ≤ 10^{-50}.

Gene	Peak score	FDR.bonf	function	P-value
MAP3776c	32.51	4.18E-228	ABC transporter ATP-binding protein	4.04E-232
MAP1134	23.71	1.61E-120	16S rRNA m5C967 methyltransferase	1.55E-124
MAP4122	22.29	2.41E-106	hypothetical protein	2.33E-110
MAP2627c	22.09	2.16E-104	hypothetical protein	2.09E-108
MAP1398	21.41	5.37E-98	hypothetical protein	5.19E-102
MAP1129	20.53	5.64E-90	lysoplasmalogenase	5.46E-94
MAP3230c	20.50	1.04E-89	AraC family transcriptional regulator	1.00E-93
rpmG	19.86	4.49E-84	50S ribosomal protein L33	4.34E-88
MAP2419	19.61	7.04E-82	membrane protein	6.81E-86
MAP2389c	19.59	1.03E-81	amidohydrolase	9.96E-86
aroA	19.51	4.62E-81	3-phosphoshikimate 1-carboxyvinyltransferase	4.47E-85
MAP2370c	19.25	7.14E-79	short-chain dehydrogenase	6.90E-83
MAP2011	19.23	1.12E-78	hypothetical protein	1.08E-82
MAP0130	19.20	1.93E-78	ATP-binding protein	1.87E-82
MAP2640c	18.84	1.86E-75	CPBP family intramembrane metalloprotease	1.80E-79
MAP2620c	18.73	1.56E-74	nitrate reductase subunit alpha	1.51E-78
MAP1360	18.70	2.47E-74	phenylalanine--tRNA ligase subunit beta	2.39E-78
MAP2969c	18.65	6.83E-74	hypothetical protein	6.61E-78
MAP0351	18.64	7.32E-74	transcriptional regulator	7.07E-78
MAP3430	18.53	6.66E-73	phosphomannomutase	6.44E-77
MAP2173c	18.46	2.42E-72	pseudo	2.34E-76
MAP2465c	18.36	1.50E-71	hypothetical protein	1.45E-75
MAP0636	18.19	3.12E-70	CPBP family intramembrane metalloprotease	3.02E-74
MAP2744c	17.74	1.03E-66	catalase-related peroxidase	9.97E-71
MAP0867c	17.61	9.82E-66	LLM class F420-dependent oxidoreductase	9.50E-70

Table 1 (cont'd)

rsmD	17.59	1.40E-65	16S rRNA (guanine(966)-N(2))-methyltransferase RsmD	1.35E-69
MAP_RS19330	17.47	1.27E-64	ANTAR domain-containing protein	1.22E-68
MAP2411	17.36	8.04E-64	pyridoxamine 5'-phosphate oxidase	7.77E-68
MAP0357	17.28	3.43E-63	membrane protein	3.32E-67
MAP2395c	17.13	4.66E-62	enoyl-CoA hydratase/isomerase family protein	4.51E-66
MAP2479	17.10	7.40E-62	potassium transporter TrkA	7.16E-66
rsgA	17.10	8.08E-62	ribosome small subunit-dependent GTPase A	7.81E-66
MAP3477	16.78	1.83E-59	pseudo	1.77E-63
MAP2747	16.72	4.92E-59	long-chain-fatty-acid--CoA ligase	4.76E-63
MAP3486	16.49	2.23E-57	lactate 2-monooxygenase	2.15E-61
MAP1560	16.49	2.40E-57	esterase	2.32E-61
MAP3063	16.45	4.56E-57	1,4-alpha-glucan-branching protein	4.41E-61
MAP3015	16.33	3.20E-56	short-chain dehydrogenase/reductase	3.10E-60
MAP1161	16.09	1.54E-54	hypothetical protein	1.49E-58
dkgA	15.88	4.35E-53	2,5-diketo-D-gluconic acid reductase	4.21E-57
MAP0988	15.82	1.18E-52	nucleoside triphosphate pyrophosphohydrolase	1.14E-56
MAP1227	15.56	6.84E-51	methylmalonyl Co-A mutase-associated GTPase MeaB	6.61E-55
MAP3488c	15.53	1.08E-50	hypothetical protein	1.04E-54

Table 2. List of genes regulated by Fur under iron deplete condition FDR ≤ 10^{-50}.

Gene	Peak score	FDR.bonf	Function	P-value
MAP_RS12480	19.48	7.85E-81	23S ribosomal RNA	7.59E-85
MAP3664	18.75	1.04E-74	glycosyl transferase	1.01E-78
rrf	18.04	4.50E-69	5S ribosomal RNA	4.35E-73
MAP3638	17.74	1.03E-66	hemophore	9.92E-71
MAP0182c	16.39	1.25E-56	hypothetical protein	1.20E-60
MAP2957	16.15	6.05E-55	peptidase M23	5.85E-59
MAP_RS12480	16.11	1.17E-54	23S ribosomal RNA	1.13E-58
MAP3471c	15.79	1.76E-52	hypothetical protein	1.70E-56
MAP_RS14585	15.51	1.61E-50	hypothetical protein	1.56E-54
MAP2961c	15.50	1.68E-50	DNA-protecting protein DprA	1.63E-54
MAP1420	15.44	4.45E-50	non-ribosomal peptide synthetase	4.30E-54

Table 3. List of genes regulated by Fur under iron replete and deplete condition FDR ≤ 10^{-50}.

		Peak score		FDR. Bonferroni	
Gene	Function	Replete	Deplete	Replete	Deplete
MAP3736c	ABC transporter ATP-binding protein	38.74	33.46	0.00	9.02E-242
MAP2381	acetoin dehydrogenase	26.50	26.02	5.07E-151	1.70E-145
MAP2071c	cyclohexanecarboxylate-CoA ligase	17.59	21.33	1.49E-65	3.10E-97
MAP2840c	diaminopimelate epimerase	19.99	16.79	3.42E-85	1.39E-59

Table 4. List of genes regulated by Fur under iron replete (yellow) and iron deplete (gray) condition with no function assigned FDR ≤ 10^{-50}.

Gene	Peak score	FDR.bonf	Binding location	Function	P-value
MAP4122	22.29	2.41E-106	intragenic	hypothetical protein	2.33E-110
MAP2627c	22.09	2.16E-104	intergenic	hypothetical protein	2.09E-108
MAP1398	21.41	5.37E-98	intergenic	hypothetical protein	5.19E-102
MAP2011	19.23	1.12E-78	intragenic	hypothetical protein	1.08E-82
MAP2969c	18.65	6.83E-74	intragenic	hypothetical protein	6.61E-78
MAP2173c	18.46	2.42E-72	intragenic	pseudo	2.34E-76
MAP2465c	18.36	1.50E-71	intragenic	hypothetical protein	1.45E-75
MAP3477	16.78	1.83E-59	intragenic	pseudo	1.77E-63
MAP1161	16.09	1.54E-54	intragenic	hypothetical protein	1.49E-58
MAP3488c	15.53	1.08E-50	intergenic	hypothetical protein	1.04E-54
MAP0182c	16.39	1.25E-56	intragenic	hypothetical protein	1.20E-60
MAP3471c	15.79	1.76E-52	intragenic	hypothetical protein	1.70E-56
MAP_RS14585	15.51	1.61E-50	intragenic	hypothetical protein	1.56E-54

Table 5. FIMO output. Most significant Fur box motif (SRYAATGAAAATSRTTWTC) derived from ChIP-seq binding in iron-replete (yellow) and deplete (gray) conditions.

Start	End	Strand	Binding sequence	Gene
4158881	4158899	+	GATAATGAAAATCGTTATC	MAP3736c
4216788	4216806	-	GTTAATGAAAATGATTATC	MAP3772c
4213755	4213773	+	CTTATTGAAAATGATTTTC	MAP3770
3880225	3880243	+	GATAACGATAATCATTTTC	MAP3490
3880225	3880243	-	GAAAATGATTATCGTTATC	MAP3489c
4213810	4213828	+	CGAGATGAAAATGATTCCC	MAP3770
4219972	4219990	-	AGATATGAAAACGGTTATC	MAP3776c
1029910	1029928	+	GACGCTGAAAGTGCTTTGC	MAP0988
3341431	3341449	-	GGTGACGAAAGTGGTTCTG	MAP3004c
2676939	2676957	+	GGTGATGACCAACGTTCCC	MAP2382
2718950	2718968	-	CACAGGGAAATTGGTCCTG	MAP2420c
3880219	3880237	+	GCTGCTGATAACGATAATC	MAP3490
4071822	4071840	+	CACCGCGAAAAGCGTTGTG	MAP3665c
2752597	2752615	+	CACACCGCAAAAGCTTTCC	MAP2451c
4158881	4158899	+	GATAATGAAAATCGTTATC	MAP3736c

REFERENCES

REFERENCES

Alexander, D.C., Turenne, C.Y., Behr, M.A. (2009) Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. J Bacteriol. 191(3):1018- 25. DOI: 10.1128/JB.01340-08

Bagg, A., and Neilands, J.B. (1987) Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. Biochemistry 26: 5471–5477.

Bannantine, J.P., Baechler, E., Zhang, Q., Li, L., Kapur, V. (2002) Genome scale comparison of *Mycobacterium avium* subsp. *paratuberculosis* Reveals potential diagnostic sequences. J Clin Microbiol. 40(4):1303-10. doi: 10.1128/JCM.40.4.1303–1310.2002

Bannantine, J.P., Paustian, M.L. (2006) Identification of Diagnostic Proteins in *Mycobacterium avium* subsp. *paratuberculosis* by a Whole Genome Analysis Approach. Diagnostic Bacteriology- Protocol – Second Edition. New Jersey, Humana Press.

Bereswill S, Greiner S, van Vliet AH, Waidner B, Fassbinder F, Schiltz E, Kusters JG, Kist M. (2000). Regulation of ferritin-mediated cytoplasmic iron storage by the ferric uptake regulator homolog (Fur) of Helicobacter pylori. J. Bacteriol. 182:5948–5953.

Browning, D.F. and Busby, S.J. (2004) The regulation of bacterial transcription initiation. Nat. Rev. Microbiol., 2, 57–65. doi: 10.1038/nrmicro787

Butcher, B.G., Bronstein, P.A., Myers, C.R., Stodghill, P.V., Bolton, J.J., Markel, E.J., Filiatrault, M.J., Swingle, B., Gaballa, A., Helmann, J.D., Schneider, D.J., Cartinhour, S.W. (2011) Characterization of the Fur Regulon in *Pseudomonas syringae* pv. tomato DC3000 Journal of Bacteriology. 193:18 -4598–4611 doi:10.1128/JB.00340-11

Chandrangsu, P., Rensing, C., Helmann, J. (2017) Metal homeostasis and resistance in bacteria. Nat Rev Microbiol. 15(6):338-350. doi: 10.1038/nrmicro.2017.15.

Collas, P. (2019) Chromatin Immunoprecipitation Assays. Methods in Molecular Biology. 567(1)1-26 doi: 10.1007/978-1-60327-414-2

Delany, I., Spohn, G., Rappuoli, R., and Scarlato, V. (2001) The Fur repressor controls transcription of iron-activated and -repressed genes in *Helicobacter pylori*. *Mol Microbiol* 42: 1297–1309

Dillon, S.C., Espinosa, E., Hokamp, K., Ussery, D.W., Casadesus, J., Dorman, C.J. (2012) LeuO is a global regulator of gene expression in *Salmonella enterica* serovar Typhimurium. Mol. Microbiol. 85:1072–1089. doi: 10.1111/1462-2920.13838

Eckelt, E., Jarek, M., Fromke, C., Meens, J., Goethe, R. (2014) Identification of a lineage specific zinc responsive genomic island in *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics. 15:1076. doi: 10.1186/1471-2164-15-1076

Escolar, L., Perez-Martin, J., de Lorenzo, V. (1999) Opening the iron box: transcriptional metalloregulation by the Fur protein. *J Bacteriol* 181: 6223–6229.

Fitzgerald, D.M., Bonocora, R.P., Wade, J.T. (2014) Comprehensive Mapping of the *Escherichia coli* Flagellar Regulatory Network. PLoS Genet. doi: 10.1371/journal.pgen.1004649

Garcia, A.B., Shalloo, L. (2015) The economic impact and control of paratuberculosis in cattle. J. Dairy Sci. 98:5019–5039. doi:10.3168/jds.2014-9241

Hantke, K. (1981) Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol Gen Genet 182: 288–292.

Hantke, K. (2001) Iron and metal regulation in bacteria. Current Opinion in Microbiology. 4:172-177.

Helmann, J.D. (2014) Specificity of Metal Sensing: Iron and Manganese Homeostasis in *Bacillus subtilis.* The Journal of Biological Chemistry. 289:41- 28112–28120. doi:10.1074/jbc.R114.587071

Holmes, K., Mulholland, F., Pearson, B.M., Pin, C., McNicholl-Kennedy, J., Ketley, J.M., Wells, J.M. (2005). *Campylobacter jejuni* gene expression in response to iron limitation and the role of Fur. Microbiology 151:243– 257. doi: 10.1099/mic.0.27412-0

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Janagama, H.K., Lamont, E.A., George, S., Bannantine, J.P., Xu, W.W., Tu, Z.J., Wells, S.J., Schefers, J., Sreevatsan, S. (2010) Primary transcriptomes of *Mycobacterium avium* subsp. *paratuberculosis* reveal proprietary pathways in tissue and macrophages. BMC Genomics. 11, 561. doi: 10.1186/1471-2164-11-561.

Karimzadeh, M., Hoffman, M.M. (2018) Virtual ChIP-seq: predicting transcription factor binding by learning from the transcriptome. bioRvix. doi: http://dx.doi.org/10.1101/168419.

Knapp, G.S., Lyubetskaya, A., Peterson, M.W., Gomes, A.L.C., Ma, Z., Galagan, J.E., McDonough, K.A. (2015) Role of intragenic binding of cAMP responsive protein (CRP) in regulation of the succinate dehydrogenase genes Rv0249c-Rv0247c in TB complex mycobacteria Nucleic Acids Research. 43:11- 5377–5393 doi:10.1093/nar/gkv420

Lamont, E.A., Xu, W.W., Sreevatsan, S. (2013) Host-Mycobacterium avium subsp. paratuberculosis interactome reveals a novel iron assimilation mechanism linked to nitric oxide stress during early infection. BMC Genomics. 14:694 doi: 10.1186/1471-2164-14-694

Li, L., Bannantine, J.P., Zhang, Q., Amonsin, A., May, B. J., Alt,D., Banerji, N., Kanjilal, S., Kapur, V. (2005) The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc. Natl. Acad. Sci. 102:12344-12349. doi:10.1073/pnas.0505662102

Lee, J.W., Helmann, J.D. (2007) Functional specialization within the fur family metalloregulators. Biometals. 20:485-499. doi: https://doi.org/10.1007/s10534-006-9070-7

Masse, E., Vanderpool, C.K., Gottesman, S. (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. J Bacteriol 187:6962–6971. doi.org/10.1128/JB.187.20.6962-6971.2005.

Miles, S., Carpenter, B.M., Gancz, H., Merrell, D.S. (2010) Helicobacter pylori apo-Fur Regulation Appears Unconserved Across Species. The Journal of Microbiology 48:3-378-386 doi: 10.1007/s12275-010-0022-0

Mills, S.A., Marletta, M.A. (2005) Metal binding characteristics and role of iron oxidation in the ferric uptake regulator from *Escherichia coli.* Biochemistry 44:13553-13559. doi: 10.1021/bi0507579

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., Jahn, D. (2005) Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics 21(22):4187-4189. doi.org/10.1093/bioinformatics/bti635

Rathnaiah, G., Zinniel, D.K., Bannantine J.P., Stabel, J.R., Gröhn, Y.T., Collins, M.T., Barletta, R.G. (2017) Pathogenesis, Molecular Genetics, and Genomics of *Mycobacterium avium* subsp. *paratuberculosis*, the etiologic Agent of Johne's Disease. Front Vet Sci. 4:187. doi:10.3389/fvets.2017.00187

Rodriguez, G.M., Voskull, M.I., Gold, B., Schoolnik, G.K., Smith, I. (2002) IdeR, an essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism, and oxidative stress response. Infect Immun. 70 (7): 3371-3381. doi: 10.1128/IAI.70.7.3371-3381.2002.

Stratmann, J., Strommenger, B., Goethe, R., Dohmann, K., Gerlach, G.F., Stevenson, K., Li, L., Zhang, Q., Kapur, V., Bull, T. (2004) A 38-kilobase pathogenicity island specific for *Mycobacterium avium* subsp. *paratuberculosis* encodes cell surface proteins expressed in the host. Infect. Immun. 72, 1265–1274. DOI: 10.1128/IAI.72.3.1265- 1274.2004

Tullius, M.V., Harmston, C.A., Owens, C.P., Chim, N., Morse, R.P., McMath, L.M., Iniguez, A., Kimmey, J.M., Sawaya, M.R., Whitelegge, J.P., Horwitz, M.A., Goulding, C.W. (2011) Discovery and characterization of a unique mycobacterial heme acquisition system. Proc Natl Acad Sci USA 108(12):5051-5056.

Wang, J., Pritchard, J.R., Kreitmann, L., Montpetit, A., Behr, M.A. (2014) Disruption of *Mycobacterium avium* subsp. *paratuberculosis*-specific gene impairs in vivo fitness. BMC Genomics. 15:415. doi: 10.1186/1471-2164-15-415

Wang, J., Moolji, J., Dufort, A., Staffa, A., Domenech, P., Reed, M.B., Behr, M.A. (2016) Iron acquisition in *Mycobacterium avium* subsp. *paratuberculosis.* J. Bacteriol. 198:857–866. DOI:10.1128/JB.00922-15

Wee S, Neilands JB, Bittner M, Hemming B, Haymore B, Seetharam R. (1988) Expression, isolation and properties of Fur (ferric uptake regulation) protein of *Escherichia coli* K12. Biol Metals. 1: 62–68

Zhu, X., Tu, Z.J., Coussens, P.M., Kapur, V., Janagama, H., Naser, S., Sreevatsan S. (2008) Transcriptional analysis of diverse strains *Mycobacterium avium* subspecies *paratuberculosis* in primary bovine monocyte derived macrophages. Microbes and Infection. 1274-1282. doi:10.1016/j.micinf.2008.07.025

CHAPTER FIVE: ARE TRANSCRIPTIONAL RESPONSES OF *MYCOBACTERIUM AVIUM* SUBSP. *PARATUBERCULOSIS* TO IRON AVAILABILITY MEDIATED BY THE FERRIC UPTAKE REGULATOR?

5.1 ABSTRACT

Mycobacterium avium subsp. *paratuberculosis* (MAP) requires supplementation of the siderophore mycobactin J for optimal growth in culture media. Recent studies characterized a Fur ortholog (*MAP3773c*) that recognizes a 19bp consensus Fur box in a metal-dependent manner. Furthermore, ChIP-seq analysis identified putative Fur binding sites on the MAP genome under iron replete and deplete conditions. In this study, we sought to determine transcriptomes of *Mycobacterium avium* subsp. *paratuberculosis.* An in-frame deletion of *MAP3773c* was accomplished using homologous recombination and genome-wide transcriptional differences were examined between parent, mutant, and complemented strains, under either iron-replete or -deplete conditions. In-frame deletion of *MAP3773c* was confirmed by PCR and sequencing. RNA samples with RNA integrity number greater than 7 were selected for RNA-Seq. Six biological replicates from each strain and condition were sequenced (n = 36 samples). Using a cutoff of 95% or more of reads aligning to the MAP K10 reference genome, 15 samples were then used for transcriptional profiling. Bioinformatic analysis showed no statistically significant differences in gene expression profiles between strains and iron conditions, but this may have been due to limited statistical power after dropping samples below the cutoff rather than a true lack of differences. While these transcriptional differences were not informative, the ChIP-seq data and the multiple networks linked to iron metabolism still suggest Fur_{MAP} likely plays a role in some cell stress responses. Additional work is needed to establish the regulation of Fur_{MAP} and understand its involvement in MAP survival and virulence.

5.2 INTRODUCTION

Bacteria need ferrous iron for most of metabolic processes and basic survival. Pathogenic mycobacteria are no exception, and they experience a variety of additional stressors inside host cells, and their ability to maintain iron homeostasis under stress represents a major virulence determinant (Rodriguez and Smith, 2003). In mycobacteria, specialized iron-acquisition systems known as mycobactins are produced by a group of genes comprised of *mbtA-mbtJ* (McMahon, Rush and Thomas, 2012; Golonka, Yeoh and Vijay-Kumar, 2019). In *Mycobacterium avium* subsp. *paratuberculosis* (MAP), the *mbt* operon is directly involved in iron metabolism and regulated by the iron-dependent repressor protein (IdeR) (Janagama *et al.,* 2009).

In addition to IdeR, multiple investigators have suggested the existence of an alternative iron uptake mechanism (Stratmann *et al.,* 2004; Eckelt *et al.,* 2014; Wang *et al.,* 2016). Located in the putative metal transport MAP-specific operon – large genomic polymorphism (LSP) 15 – is *MAP3773c* (Alexander *et. al*., 2009). In a recent study, we cloned, purified and performed full characterization of *MAP3773c,* confirming that it encodes a ferric uptake regulator (Fur) ortholog that recognizes a 19-bp DNA sequence motif (Fur box), and that it is involved in metal homeostasis (Shoyama *et al*., 2020). The transcriptional regulator Fur is well-known in other bacteria for controlling iron metabolism, regulating defenses against oxidative stress, and is considered the master regulator of iron homeostasis (Hantke, 1981; Ma, Faulkner and Helmann, 2012). In general, in an iron-replete environment, Fur binds to intracellular Fe^{2+} and forms the Fur–Fe^{2+} complex, transitioning Fur from its inactive (apo-) to its activated (holo-) form that recognizes the Fur box sequence (Hantke, 2001). This process results in the transcription repression of most its regulon by blocking the entry of RNA polymerase to the promoter region (Hantke, 2001). However, some studies have shown that Fur can also function as an activator of

gene expression (Pi and Helmann, 2018). Through an indirect mechanism mediated by a small regulatory RNA known as RhyB, Fur was demonstrated to positively regulate a set of genes that encode iron-containing proteins in *E. coli,* including ferritins, a superoxide dismutase, and some genes in the tricarboxylic acid (TCA) cycle (Masse and Gottesman, 2002). Furthermore, in *Neisseria meningitidis*, investigators showed that when Fur binds to operators located upstream of three promoters regulated by Fur and iron, it results in activation of transcription (Delany, Rappuoli and Scarlato, 2004). In contrast, when the binding site overlaps the promoter region, Fur triggers repression (Delany, Rappuoli and Scarlato, 2004).

Whether Fur in *Mycobacterium avium* subsp. *paratuberculosis* (MAP) acts as an activator or repressor is still unknown. ChIP-seq analysis identified several putative binding sites and genes involved in iron homeostasis in MAP; in addition, MAP Fur was confirmed as a global regulator that recognizes many target sites in the genome, either under iron-replete or deplete conditions (Shoyama *et al*., 2020). In this study, we sought to determine transcription alterations regulated by Fur_{MAP} in response to iron limitation by RNA-Seq. We constructed an in-frame *Fur* mutant of MAP strain K10 through homologous recombination, and a complementation strain of this mutant background transformed with a *Fur*-bearing plasmid to help identify genes that were exclusively regulated by Fur.

5.3 MATERIALS AND METHODS

5.3.1 BACTERIAL CULTURE AND CONDITIONS

Mycobacterium avium subsp. *paratuberculosis* K10, mutant *MAP3773c (ΔMAP3773c*) and complemented strains were grown at 37°C shaking in Middlebrook 7H9 supplemented with 10% OADC (oleic acid, dextrose, catalase) enrichment medium (Thermo Fisher Scientific, Waltham, MA, United States), 0.05% Tween 80, and 2mg/L of ferric mycobactin J (Allied Monitor Inc.,

Fayette, MO, United States). Antibiotics – 50µg/mL kanamycin, or 150µg/mL hygromycin – were added when necessary. *E. coli* TOP10F cells (Invitrogen, Carlsbad, CA, United States) were grown in LB medium 37°C with shaking at 200 RPM.

5.3.2 *MAP3773C* MUTANT

The DNA sequence of *MAP3773c* was downloaded from the KEGG MAP genome site (http://www.genome.jp/dbget-bin/www_bget?mpa:MAP_3773c) and primers were designed to amplify the upstream region including 84bp of *MAP3773c* (*MAP3773c* UPS fwd: GCCTCGGTACCTGCCTCGGTCAATCCGGTAG; rev: CTCTCTAGAGTTCTCTTGCGCTCGCAGCAC) and a second set for the downstream region including 111bp of the *MAP3773c* coding C-terminus region (*MAP3773c* DWN fwd: CTCAAGCTTCGGCTAAGCCGCCAACATCA; REV: CTCCTCGAGCAAGTAGGTCGGCAATCGTG).

The PCR product of *MAP3773c* UPS was cloned into the cosmid vector pYUB854 carrying hygromycin resistance (Bardarov *et al.*, 2002). The PCR product of *MAP3773c* DWN was cloned into the resulting plasmid *MAP3773c* UPS+pYUB854, generating the recombinant cosmid pBUN445.

This plasmid was cloned into the conditionally replicating shuttle phasmid vector phAE87. pBUN445+phAE87 was transfected to *M. smegmatis* MC^2 155 and plated for phage plaques (mycobacteriophage) at 30°C. Mycobacteriophage were electroporated into MAP cells and plated on selective media containing 150µg/mL hygromycin and incubated at 37°C.

5.3.3 CONSTRUCTION OF MAP COMPLEMENTATION OF MUTANT STRAIN

Complementation of the *MAP3773c* mutant was accomplished by cloning and transformation of the mutant with *MAP3773c* on a plasmid (pSM417). The *MAP3773c* gene was amplified using 20pmol of each primer (fwd: GTGAACCTGATCGCCAAGAT, rev: CTCGGTCTGCGTGTTGTAGA). PCR was carried out by standard methods with 25ng MAP K10 genomic DNA, HotStarTaq Master Mix (Qiagen, Germantown, MD, United States). PCR product was purified using PCR purification kit (Qiagen, Germantown, MD, United States). Sanger sequencing to confirm amplification of *MAP3773c* was performed at Michigan State University Genomics Core. The constitutive plasmid pSM417 with *hsp60* promoter was used as a vector (Kugadas et *al*., 2016). Double digestion of *MAP3773c* and pSM417 was performed using enzymes PstI-HF and BamHI-HF (New England Biolabs, Ipswich, MA, United States) followed by quick ligation protocol (New England Biolabs). The ligation product *MAP3776c*-psM417 was added into the chemically competent *E. coli* TOP10 (Invitrogen, Carlsbald, CA) and single colonies were selected for Sanger sequencing.

Electroporation was carried out following protocol described by Goude & Parish, 2008. A total of 0.5µg purified *MAP3773c*-pSM417 DNA was added into 0.2mL of electrocompetent *MAP3776c::Tn* cells for electroporation and subject to a single pulse of 2.5kV, 25µF, with the pulse-controller resistance set at 1000Ω. Cells were recovered with 10mL of Middlebrook 7H9 supplemented with 10% OADC (oleic acid, dextrose, catalase) enrichment medium (Thermo Fisher Scientific, Waltham, MA, United States) (M7H9-OADC) and after 16h incubation at 37°C, cells were plated on solid M7H9-OADC containing kanamycin (50µg/mL) and hygromycin (100µg/mL) until colonies became visible.

Single colonies were transferred to liquid M7H9-OADC with kanamycin (50µg/mL) and hygromycin (100µg/mL) and incubated at 37°C until OD_{600} of ~0.5 was reached. Successful integration of *MAP3773c* was confirmed by PCR and sequencing.

5.3.4 PREPARATION OF ELECTROCOMPETENT MYCOBACTERIA

To obtain an electrocompetent *MAP3773c* mutant, procedures were carried out following protocols described by Goude & Parish, 2008. In summary, a single colony was used to inoculate 10mL of M7H9-OADC medium containing kanamycin (50µg/mL) and hygromycin (100µg/mL) and incubated at 37°C until OD_{600} of ~0.5 was reached. Culture was combined with 100mL of M7H9-OADC medium and antibiotics until OD_{600} of ~0.5 was reached. The day before cells were harvested, glycine was added to a final concentration of 2M. The following day, cells were washed with gradually decreasing volumes (50, 25, 10, 5mL) of 10% glycerol and 220µL was aliquoted and frozen at -80°C.

5.3.5 RNA EXTRACTION

RNA extraction of each strain – MAP K10, *MAP3773c* mutant and complemented strain – was performed following a protocol developed by Rhode *et al.* (2007) with modifications. Five mL of bacterial culture with OD_{600} of ~0.5 was used. Iron starvation was accomplished by treating cultures with 2,2'-bipyridyl (DIP, 200µM final) for 2 hours shaking at 200rpm at 37°C (Thompson *et al.,* 2002; Eckelt *et al.,* 2014). Cells from iron-replete and -deplete conditions were washed with 1x PBS. GTC buffer (4M guanidine thiocyanate, 0.5% sodium N-lauryl sarcosine, 25mM sodium citrate (pH 7.00) and 0.1M beta-mercaptoethanol) was added to each sample followed by centrifugation at room temperature for 3500 x *g* for 10 min. Cells were resuspended

in 1mL of GTC buffer and transferred to a 1.7mL microcentrifuge tube. Cells were washed two times with 1mL PBS + 0.1% Tween80 and centrifuged at 12,500 rpm for 5 min at room temperature. Two-hundred µL of fresh 5mg/mL lysozyme was added to the bacterial pellet and incubated for 15 min at room temperature for cell lysis. Next, 750µL of Trizol reagent at 65°C was added (Life Technologies, Carlsbad, CA, United States) and transferred to a 2mL screw-cap tube with O-ring containing 0.5mL of 0.1mm glass beads, followed by complete cell lysis with MagNA Lyser (Roche Diagnostic, Sandhofer, Germany) for 2 min at max speed. After 2 min on ice, 200µL of chloroform was added and mixed by inverting the tube. Samples were centrifuged at 12,000 rpm for 15 min at room temperature. The aqueous phase was added to a new microcentrifuge tube containing 500µL of 100% RNase-free EtOH. A 700µL volume of sample was processed by Qiagen RNeasy kit DNase treatment with the Ambion TurboDNA-free kit (Thermo Fisher Scientific, Waltham, MA, United States) according to manufacturer's instructions. To ensure RNA integrity, each sample was assessed using the Agilent RNA 6000 Nano kit with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Only RNA samples with an RNA integrity number (RIN) greater than 7 (an indication of high-quality RNA with little degradation) were used for RNA-Seq.

5.3.6 RNA-SEQUENCING AND ANALYSIS

Six biological replicates from each strain (MAP K10, *ΔMAP3773c* and complemented) in iron-replete and -deplete conditions were submitted for sequencing. Approximately 1µg of total RNA was used for library creation. The RNA-Seq library was prepared using the mRNA Seq Single Read Library preparation kit per manufacturer's instructions (Illumina Inc., San Diego, CA). cDNA libraries were sequenced using Illumina sequencing technology (Illumina HiSeq 2500) in

five lanes of the sequencing flow cell. Libraries were prepared and sequenced at the Iowa State University DNA Facility (Ames, IA).

Computational analysis was carried out on MSU's High Performance Computing Center (HPCC) clusters. FASTQ files returned after sequencing were analyzed first by FastQC v0.11.7 to ensure quality scores were within acceptable parameters (Andrew, 2018). Trimmomatic v0.36 was used to trim adapters as well as any low-quality segments of reads (Bolger, Lohse and Usadel, 2014). Cleaned reads were processed using EDGE-pro v1.3.1, with necessary input files generated from the .gb file of NCBI reference sequence NC_002944.2 using the gb2ptt Perl script (Givan, 2009; Magoc, Wood and Salzberg, 2013).

For differential expression analysis in R, output .rpkm files were initially processed by the edgeToDeseq.perl script included in the EDGE-pro package for use in the DESeq2 package (Love, Huber and Anders, 2014). Before analysis in DESeq2, the individual files for each run were checked for percent alignment to the MAP K10 genome, and samples with fewer than 95% of reads aligning by the Bowtie2-based aligner in EDGE-Pro were presumed contaminated and excluded.

All runs were also double-checked by alternate alignment with the HISAT2 program, but all output confirmed Bowtie2-based aligning and supported the likely contamination of samples (Kim, Paggi and Park, 2019). DESeq2 was performed on the resulting dataset in R 4.0.0.

5.4 RESULTS

5.4.1 GENERATION OF *MAP3773C* DELETION BY ALLELIC EXCHANGE

Using the specialized transducing mycobacteriophage developed by Bardarov *et al.* (2002), our group was able to successfully disrupt *MAP3773c* with the hygromycin reporter gene. Allelic

exchange is possible as a result of a double crossover between the homologous DNA arms flanking the disrupted gene (Bardarov *et al., 2002)*. Analysis by PCR using locus-specific primers (Table 6) confirmed the configuration of the Δ*MAP3773c* deletion allele (Fig. 23). A total of eight MAP K10 transductants carrying pBUN445+phAE87 showed bands indicating amplification of inner and outer sequences flanking *MAP3773c* (Fig. 24), an absence of the *MAP3773c* sequence itself (Fig. 25, lane 6), and the presence of hygromycin gene in its place (Fig. 25, lane 3), as would be expected for our desired recombination event.

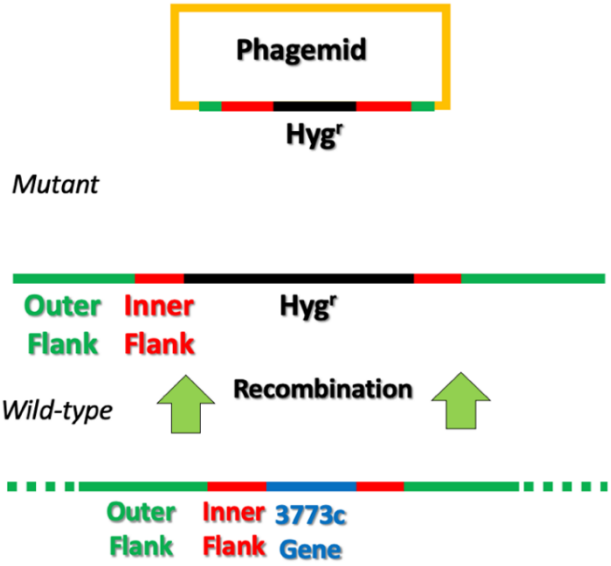


Figure 23. **Generation of MAP3773c mutant, diagram of *MAP3773c* deletion by allelic exchange.**

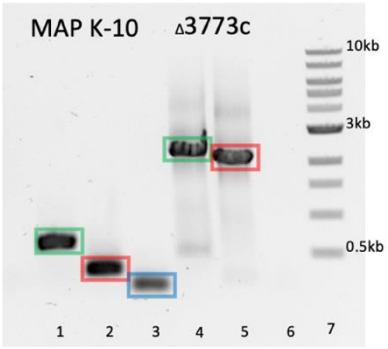


Figure 24. **0.8% Agarose gel of the three fragments of MAP3773c ORF confirming deletion generated by homologous recombination**. As indicated in Fig. 23, each color represents

Figure 24 (cont'd)
individual fragments: blue: internal *MAP3773c* replaced by the hygromycin cassette (176bp); red: inner flank (314bp for MAP K10, 1989bp for mutant) and green: outer flank (553bp for MAP K10 and 2228bp for mutant). MAP K10 DNA was used as template (lane 1-3); *MAP3773c* mutant DNA was used as template (lane 4-6). No PCR product amplification as this region was replaced by hygromycin cassette (lane 6).

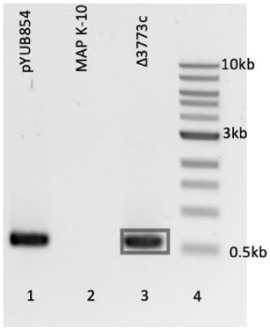


Figure 25. **0.8% Agarose gel of the three fragments of *Hygromycin* ORF confirming deletion generated by homologous recombination.** 0.8% agarose gel : Lane 1. PCR amplification of Hygromycin cassette (560 bp) from pYUB854- empty vector; Lane 2. Absence of PCR amplification for MAP K10; Lane 3. PCR amplification of pYUB854-*MAP3773c* (560 bp); Lane 4. 1kb ladder.

Table 6. PCR primers used to confirm *MAP3773c* deletion.

Primers	**5' to 3' Sequence**
Hygro-Fwd	AACACCTCGAAGTCGTGCAGGAAG
Hygro-Rev	CTACCTGGTGATGAGCCGGATGAC
MAP3773c outer flank - Fwd	CTCGGATCCAACGCGGAAGTGACGAGTAA
MAP3773c outer flank - Rev	CTCAAGCTTAGAACTCTGTGGCGGCATAA
MAP3773c inner flank - Fwd	AGCGGACCGTCTTAGAGGTG
MAP3773c inner flank - Rev	CGTCGGCGTAGTGATGTTGG
Internal *MAP3773c* gene	TCTTCGATGTCGACGGGTGTAAAGG
Internal *MAP3773c* gene	CCTGACCAGTGTTTACCGCATCTTG

5.4.2 GENOME-WIDE TRANSCRIPTOME ANALYSIS

Genome-wide expression analysis using RNA-Seq was performed to elucidate the role of Fur in iron homeostasis in *Mycobacterium avium* subsp. *paratuberculosis*. The *Fur* mutant, parent and

complemented strain were investigated under iron-replete and -deplete conditions for gene expression patterns.

FastQC reported values in the expected range for high-quality sequencing, supporting our use of the data for analysis. After alignment through the EDGE-pro and HISAT2 packages, low alignment values were reported for 21 out of the 36 samples sequenced, with some as low as ~16% of reads aligning to the MAP K10 genome. With a cutoff of 95%+ of reads aligning to MAP K10 (Fig. 26), the following 15 samples were used in the analysis (sample type/number of replicates): Wild-type replete/3; Wild-type deplete/3; Mutant replete/2; Mutant deplete/2; Complement replete/3; Complement deplete/2.

Previously, Fur binding sites in MAP have been identified by *in vivo* ChIP-seq assay; however, quantifying expression level changes in genes controlled by Fur has not been performed. For simulating identical conditions as those used in ChIP-seq for consistency, iron starvation was achieved by incubating MAP strains with the cell membrane permeable iron chelator 2,2-bipyridyl (DIP) for two hours (Shoyama *et al.,* 2020). We analyzed genes associated with binding of Fur_{MAP} and directly related to iron regulation: *MAP3776c* (ABC transporter), *MAP3638* (hemophore-like protein), *MAP3736c* (Fur box) and *MAP3837* (IdeR) (Shoyama *et al.,* 2020). To determine relationships between Fur binding and transcript levels, we compared transcript levels in MAP K10, Δ*MAP3773c* and complemented strains, under iron-replete and -deplete conditions relative to the control group (MAP K10 iron replete), but no statistically significant differences in any iron-related genes could be observed (Fig. 27).

Corroborating these findings, after excluding runs with presumptive contamination, the DESeq2 R package showed no significant differences between any of our sample types when stratified by strain (WT, Mutant, Complemented) or treatment (Iron Replete/Iron Deplete) with adjusted P

values >> 0.05. A heatmap of sample-to-sample differences shows no obvious patterns between similar or different sample types or treatments (Fig. 28). Complementing this result, principle component analysis similarly shows random distribution of all samples with no distinct clustering visible (Fig. 29).

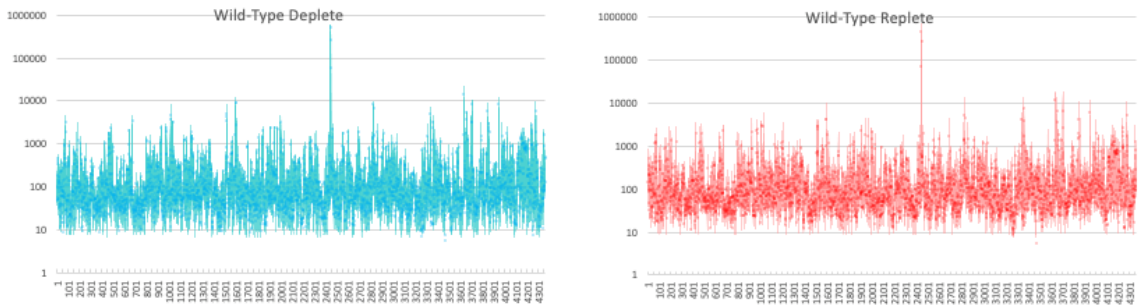


Figure 26. **Representation of alignment to the MAP K10 genome**. Y-axis: log-transformed RPKM values; X-axis: sequential gene number across the genome. Circles are mean RPKM values for each gene, lines in the background are the values for all the replicates for this sample type.

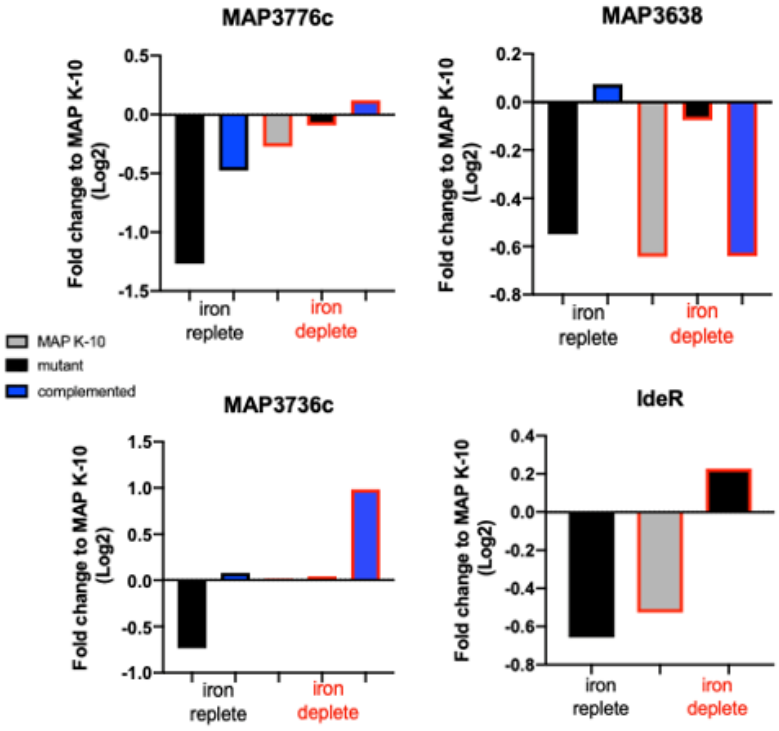


Figure 27. **Transcriptional analysis by RNA-Seq with a cutoff of 95%+ of reads aligning to MAP K10.** Gene expression of *MAP3776c, MAP3638*, *MAP3736*c and *MAP2827* (IdeR) in

Figure 27 (cont'd)
different strains and conditions were compared to MAP K10 under iron replete conditions. Bars represent the log2 fold change when compared to control (MAP K10 under iron replete condition), analyzed by DESeq2. Each bar color represents a different strain: MAP K10 (grey), *MAP3773c* mutant (black) and complemented strain (blue). Border color represents iron conditions, replete (black) and deplete (red).

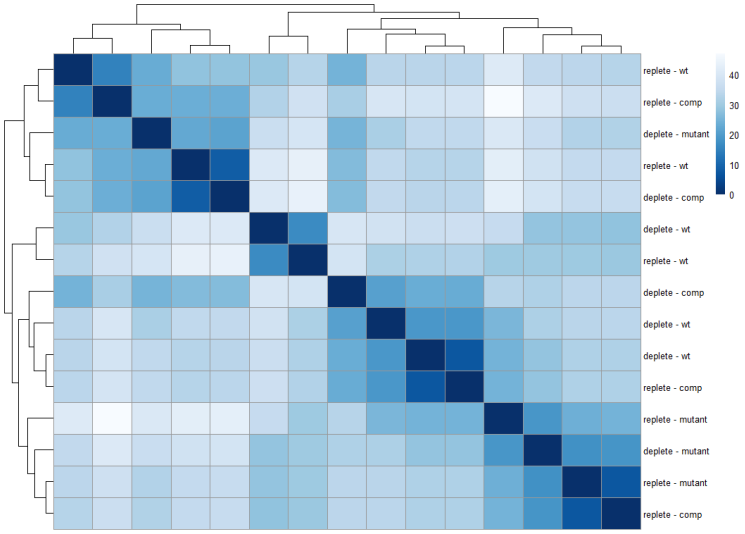


Figure 28. **Heatmap of sample-to-sample distance of transformed count matrix**. Distance-based hierarchical clustering is seen on the left and top of the figure with the labeling for the clustering on the right side of the figure. Each box in the grid represents the distance between the intersecting samples, with the darkest color representing no difference (e.g., the diagonal line comparing self-to-self samples) and lighter colors indicating greater distance. No trend is observed in the distances between and within sample groups. Generated using the deseq2, pheatmap, and rColorBrewer R packages.

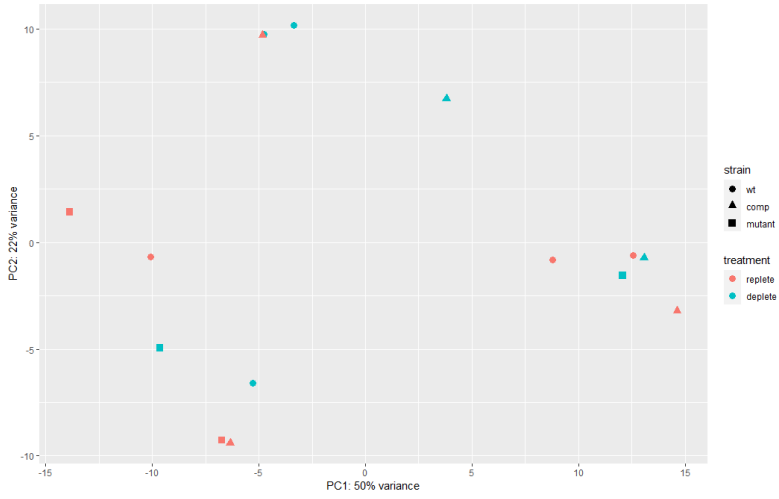


Figure 29. **Principal component analysis plot of replicates by sample type**. The distance on the X and Y axis represent the maximum data variance observable by created uncorrelated variables.

Figure 29 (cont'd)
Salmon-colored points indicate iron replete samples, while cyan-colored points indicate iron deplete samples. The shape of each point indicates the genetic background of the sample. Sample types appear dispersed randomly across ~72% of variance with no obvious patterns correlating either treatment or genetic background to the variance displayed, suggesting no differences exist between these samples at a transcriptome-wide level. Created with deseq2 and ggplot2 R packages.

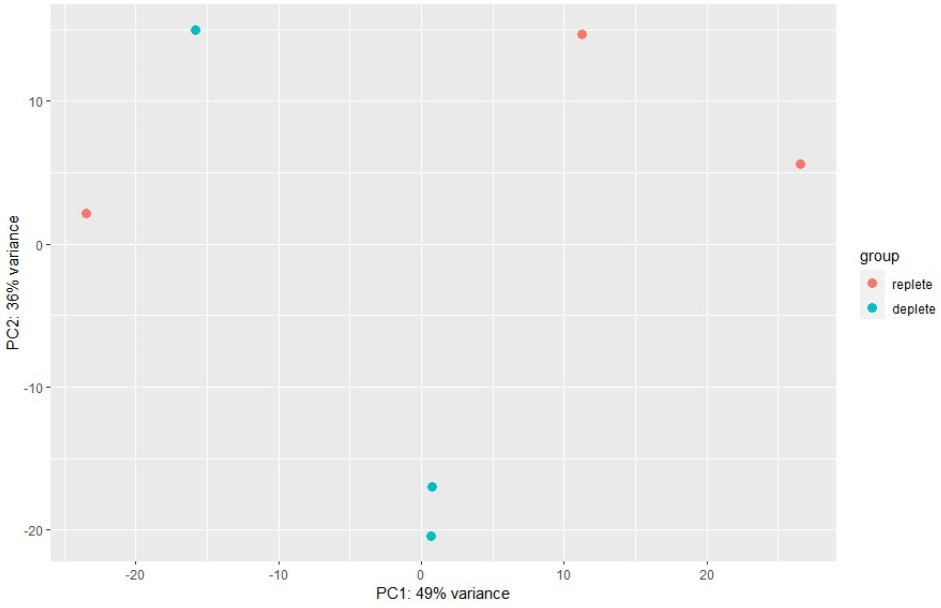


Figure 30. **Principal component analysis plot of only wild-type replicates.** The distance on the X and Y axis represent the maximum data variance observable by created uncorrelated variables. Salmon-colored circles indicate iron replete samples, while cyan-colored circles indicate iron deplete samples. Sample types appear dispersed randomly across ~85% of variance with no obvious patterns correlating either treatment or genetic background to the variance displayed, suggesting no differences exist between these samples at a transcriptome-wide level. Created with deseq2 and ggplot2 R packages.

c_d_2_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Uncultured bacterium	183	37.89
Pelosinus sp.	103	21.33
Pelosinus fermentans	97	20.08
Methylomusa anaerophila	86	17.81
Uncultured Propionispora	81	16.77
Propionispora hippei	76	15.73
uncultured Propionispora	54	11.18
Propionispora vibrioides	48	9.94
Veillonella rodentium	34	7.04

c_d_5_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	483	61.29
Paenibacillus sp.	371	47.08
Paenibacillus lautus	289	36.68
Uncultured bacterium	126	15.99
Paenibacillus uliginis	112	14.21
Paenibacillus sabinae	80	10.15
Paenibacillus durus	68	8.63
Paenibacillus riograndensis	66	8.38
Paenibacillus kribbensis	61	7.74

c_r_1_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	506	67.02
Paenibacillus sp.	379	50.20
Paenibacillus lautus	310	41.06
Paenibacillus uliginis	78	10.33
Uncultured bacterium	78	10.33
Paenibacillus sabinae	76	10.07
Paenibacillus durus	58	7.68
Paenibacillus riograndensis	56	7.42
Paenibacillus graminis	41	5.43

c_r_3_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	373	54.45
Paenibacillus sp.	273	39.85
Paenibacillus lautus	218	31.82
Uncultured bacterium	119	17.37
Paenibacillus uliginis	86	12.55
Paenibacillus sabinae	50	7.30
Pelosinus sp.	45	6.57
Methylomusa anaerophila	41	5.99
Pelosinus fermentans	39	5.69

m_d_3_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	532	70.00
Paenibacillus sp.	400	52.63
Paenibacillus lautus	326	42.89
Paenibacillus uliginis	103	13.55
Uncultured bacterium	88	11.58
Paenibacillus sabinae	81	10.66
Paenibacillus riograndensis	69	9.08
Paenibacillus durus	66	8.68
Paenibacillus graminis	55	7.24

m_r_1_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	618	72.62
Paenibacillus sp.	488	57.34
Paenibacillus lautus	397	46.65
Paenibacillus uliginis	179	21.03
Uncultured bacterium	153	17.98
Paenibacillus sabinae	103	12.10
Paenibacillus riograndensis	94	11.05
Paenibacillus durus	90	10.58
Paenibacillus graminis	77	9.05

m_r_3_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	308	48.89
Paenibacillus sp.	245	38.89
Paenibacillus lautus	197	31.27
Uncultured bacterium	90	14.29
Paenibacillus uliginis	69	10.95
Pelosinus fermentans	58	9.21
Pelosinus sp.	54	8.57
Methylomusa anaerophila	45	7.14
Paenibacillus sabinae	44	6.98

w_d_3_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	504	66.58
Paenibacillus sp.	385	50.86
Paenibacillus lautus	309	40.82
Paenibacillus uliginis	111	14.66
Uncultured bacterium	95	12.55
Paenibacillus sabinae	76	10.04
Paenibacillus riograndensis	61	8.06
Paenibacillus durus	57	7.53
Paenibacillus kribbensis	41	5.42

w_d_4_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Homo sapiens	159	26.07
Paenibacillus glucanolyticus	95	15.57
Human DNA	74	12.13
Cutibacterium acnes	67	10.98
Paenibacillus sp.	66	10.82
Propionibacterium acnes	65	10.66
Uncultured bacterium	55	9.02
Paenibacillus lautus	53	8.69
Eukaryotic synthetic	51	8.36

w_r_1_unmapped_1k_top_report.txt

Organism	reads_assigned	pct
Paenibacillus glucanolyticus	450	60.32
Paenibacillus sp.	346	46.38
Paenibacillus lautus	258	34.58
Uncultured bacterium	144	19.30
Paenibacillus uliginis	86	11.53
Uncultured Paenibacillus	64	8.58
Micromonospora auratinigra	48	6.43
Mycobacterium sp.	43	5.76
Bacillus atrophaeus	42	5.63

Figure 31. **Meta-analysis of the top 10 unaligned reads by BLAST.**

5.5 DISCUSSION

High-throughput technologies allow researchers to investigate fine biological aspects of cell function and better understand diseases etiologies. In this study, we sought to elucidate iron regulation in MAP by integrating previously identified Fur_{MAP} binding sites to RNA-Seq. However, a comparative analysis of the transcriptomes from wild-type MAP K10, a *Fur* deletion mutant, and a complemented strain showed no significant effects at the transcriptional level under the studied conditions, despite Fur's known MAP genome-binding as shown by ChIP-seq assays (Shoyama *et al.*, 2020). The PCA analysis and distance matrix of our samples cast doubt on meaningful differences between sample types at a broader level (Fig. 29), and narrowing the analysis to only wild-type controls in iron-replete or iron-depleted conditions still shows a random clustering pattern (Fig. 30). Given that no repeatable differences were observed even

between wild-type controls, the lack of differential gene expression may not reflect biological reality. Possible explanations to this unexpected result may lie with the experimental design, and likely contamination of several replicates and subsequent exclusion from analysis severely restricted our statistical power.

Preparation of RNA samples for sequencing is a multistep process involving reverse transcription, amplification, fragmentation, purification, and adaptor ligation (Auer and Doerge, 2010). In addition to these manipulations, next generation sequencing (NGS) is exquisitely sensitive as a molecular biology method, and one of the main challenges to high-throughput RNA-Seq is contamination (Nieuwenhuis *et al.,* 2020). In our study, 21 out of 36 samples showed low alignment values when mapped against the MAP K10 reference genome, an indication of contamination that was supported by the meta-analysis by BLAST of reads that failed to align to the K10 genome (Fig. 31). Various types of contamination in sequencing runs have been reported in other studies – contamination has been seen in different points along the sequencing process, from the research laboratory where samples are being extracted, to the sequencing facility where libraries are being prepared and handled (Hadfield and Eldridge, 2014; Sangiovanni *et al.,* 2019). Although quality control checks were done before sequencing and all samples were prepared in a DNA-free environment, treated with DNAse, and a run through a subtractive hybridization procedure by sample treatment with RiboMinus to deplete ribosomal RNA and enrich coding sequences, we still had 21 samples contaminated with exogenous material. Further investigation identifying these contaminating species can help imply at which sample preparation or library construction step it might be introduced into the sample, and thus give instruction on avoiding similar accidents in the future (Zhou *et al.,* 2018).

Another additional challenge of high-throughput RNA-Seq is the variation within experiments. Besides processing date, technician, and reagent batch, other technical effects such as library preparation and technology-specific effects during the sequencing process can affect the final result (Fang and Cui, 2011; Nieuwenhuis *et al.,* 2020). To overcome the limitations and improve the quality and reliability of differential expression (DE) detection by RNA-Seq, studies suggest that a low-replicate experiment may not provide a sufficiently detailed view of the DE to represent biology accurately; thus, multiple biological replicates are not only desirable but necessary (Fang and Cui, 2011; Robles *et al.,* 2012). When we initially incorporated 6 biological replicates from each strain/treatment, our goal was to have enough power to detect DE. Instead, we ended up with fewer replicates (Wild-type replete/3; Wild-type deplete/3; Mutant replete/2; Mutant deplete/2; Complement replete/3; Complement deplete/2) that yielded potentially inadequate statistical power (Auer and Doerge, 2010; Fang and Cui 2011; Liu *et al*, 2014; Lamarre *et al.,* 2018). Furthermore, as transcription is a highly dynamic biological process, performing a time-course RNA-Seq provides an opportunity to evaluate gene expression specific to different stages of growth and response after treatment (Thomas *et al.,* 2018). Even though two hours after depletion was an informative time point for ChIP-seq, transcriptome responses could have occurred within minutes after depletion or days as shown in other studies using the same chelator (Brinkworth *et al.,* 2018; Pandey *et al.,* 2018).

5.6 CONCLUSIONS AND FUTURE DIRECTIONS

Mycobacterium avium subsp. *paratuberculosis* and its special relationship with iron is well-known, and combined with the complex iron homeostasis problems that pathogenic bacteria face from hosts, the importance of a thorough understanding of iron metabolism is reinforced.

As the current RNA-Seq analysis were inconclusive due to limitations discussed above, future studies first will need to investigate the source of contamination during RNA processing and identify the optimal time for RNA extraction after iron depletion.

In addition, future investigations with the *MAP3773c* mutant and complemented strains should involve function characterization to identify if deletion on *MAP3773c* has any effect in MAP growth and if the complemented strains was able to restore function.

REFERENCES

REFERENCES

Alexander, D.C., Turenne, C.Y., Behr, M.A. (2009) Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. J Bacteriol. 191(3):1018- 25. DOI: 10.1128/JB.01340-08

Andrews, S. (2018). FASTQC: A quality control tool for high throughput sequence data [v.0.11.7].

Auer, P.L., Doerge, R.W. (2010). Statistical Design and Analysis of RNA sequencing Data. Genetics. 185(2):405-416. Doi:10.1534/genetics.110.114983.

Bardarov, A., Bardarov Jr, S., Pavelka Jr, M.S., Sambandamurthy, V., Larsen, M., Tufariello, J., Chan, J., Hatfull, G., Jacobs, W.R. (2002). Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in *Mycobacterium tuberculosis, M. bovis BCG and M. smegmatis.* Microbiology. (148), 3007-3017. Doi:10.1099/00221287-148-10-3007.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, btu170.

Delany, I., Rappuoli, R., Scarlato, V. (2004). Fur Functions as an activator and as a repressor of putative virulence genes in Neisseria meningitidis. Mol. Microbiol. (52) 1081-1090. Doi:101111/j.13652958.2004.04030.x

Eckelt, E., Jarek, M., Fromke, C., Meens, J., Goethe, R. (2014) Identification of a lineage specific zinc responsive genomic island in *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics. 15:1076. doi: 10.1186/1471-2164-15-1076

Escolar, L., Pérez-Martín, J., de Lorenzo, V. (1999). Opening the iron box: transcriptional metalloregulation by the Fur protein. J. Bacteriol. 181, 6223-9.

Fang, Z., Cui, X. (2011). Design and validation issues in RNA-seq experiments. Briefings in Bioinformatics. 12, 280-287. DOI:10.1093/bib/bbr004.

Fillat, M. (2014). The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. Archives of Bioch. And Byophysics. 546, 41-52.

Givan, S. (2019). gb2ptt. https://github.com/sgivan/gb2ptt.

Golonka, R., Yeoh, B.S., Vijay-Kumar, M. (2019). The iron Tug-of-War between bacterial siderophores and Innate Immunity. J Innate immune. 11, 249-262. DOI:10.1159/000494627

Hantke, K. (1981). Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. Mol Gen Genet. 182, 288-92. DOI:10.1007/bf00269672

Hantke, K. (2001) Iron and metal regulation in bacteria. Current Opinion in Microbiology. 4:172-177.

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Kim, D., Paggi, J.M., Park, C. *et al.* Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37, 907–915 (2019).

Kolde, R. (2015). pheatmap: Pretty heatmaps [Software].

Lamarre, S., Frasse, P., Zouine, M. (2018). Optimization of an RNA-Seq Differential Gene Expression Analysis Depending on Biological Replicate Number and Library Size. Front Plant Sci. Doi:10.3389/fpls.2018.00108

Love MI, Huber W, Anders S (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." *Genome Biology*, 15, 550. doi: 10.1186/s13059-014-0550-8.

Liu, Y., Zhou, j., White, K.P. (2014) RNA-seq differential expression studies: more sequence or more replication? Bioinformatics. (30) 301–304 Doi: 10.1093/bioinformatics/btt688

MA, Z., Faulkner, M.J., Helmann, J.D. (2012). Origins of specificity and cross-talk in metal ion sensing by *Bacillus subtilis* Fur. Mol Microbiol. 86, 1144-55. DOI:10.1111/mmi.12049

Magoc, T., Wood, D., Salzberg, S.L. (2013). EDGE-pro: Estimated Degree of Gene Expression in Prokaryotic Genomes. Evolutionary Bioinformatics (9)127-136.

Masse, E., Gottesman, S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli.* PNAS. (7) 4620-4625. doi:10.1073/pnas.032066599

McHugh, J.P., Rodriguez-Quinones, F., Abdul-Tehrani, H., Svistunenko, D.A., Poole, R.K., Cooper, C.E., Andrews, S.C. (2003). Global Iron-dependent Gene Regulation in Escherichia coli: A new mechanism for iron homeostasis. Journal of Biol Chem. (32) 29478-29486. Doi:10.1074/jbc.M303381200

McMahon, M.D., Rush, J.S., Thomas, M.G. (2012). Analyses of MbtB, MbtE, and MbtF Suggest Revisions to the Mycobactin Biosynthesis Pathway in *Mycobacterium tuberculosis.* Journal of Bacteriology. 194, 2809-2818. DOI:10.1128/JB.00088-12

Nieuwenhuis, T.O., Yang, S.Y., Verma, R.X. et al. (2020). Consistent RNA sequencing contamination in GTEx and other data sets. Nat Commun 11, 1933. doi:10.1038/s41467-020-15821-9

Neuwirth, E. (2018). RColorBrewer: ColorBrewer palettes [Software].

Pi, H., Helmann, J.D. (2018). Genome-Wide Characterization of the Fur Regulatory Network Reveals a Link Between Catechol Degradation and Bacillibactin Metabolism in Bacillus Subtilis. mBio. (5)1451-18. Doi:10.1128/mBIO.01451-18

Rapaport, F., Khanin, R., Liang, Y. et al. Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome Biol 14, 3158 (2013). https://doi.org/10.1186/gb-2013-14-9-r95

Rhode, K.H., Abramovitch, R.B., Russel, D.G. (2007). Mycobacterium tuberculosis Invasion of Macrophages: Linking Bacterial Gene Expression to Environmental Cues. Cell Host & Microbe. (5) 352-364. Doi:10.1016/j.chom.2007.09.006.

Robles, J.A., Qureshi, S.E., Stephen, S.J., Wilson, S.R., Burden, C.J., Taylor, J.M. (2012). Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing. BMC Genomics 484. DOI: 10.1186/1471-2164-13-484

Rodriguez, G.M., Smith, I. (2003). Mechanisms of iron regulation in mycobacteria: role in physiology and virulence. Mol. Microbiol. 1365-2958. Doi:10.1046/j.1365-2958.2003.03384.x

Shoyama, F.M., Janetanakit, T., Bannantine, J.P., Barletta, R.G., Sreevatsan, S. (2020). Elucidating the Regulon of a Fur-like Protein in *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Doi:10.3398/fmicb.2020.00598

Stratmann, J., Strommenger, B., Goethe, R., Dohmann, K., Gerlach, G.F., Stevenson, K., Li, L., Zhang, Q., Kapur, V., Bull, T. (2004) A 38-kilobase pathogenicity island specific for *Mycobacterium avium* subsp. *paratuberculosis* encodes cell surface proteins expressed in the host. Infect. Immun. 72, 1265–1274. DOI: 10.1128/IAI.72.3.1265-1274.2004

Thompson, D.K., Beliaev, A.S., Giometti, C.S., Tollaksen, S.L., Khare, T., Lies, D.P., Nealson, K.H., Lim, H., Yates III, J., Bradt, C.C., Tiedje, J.M., Zhou, J. (2002). Transcriptional and Proteomic Analysis of a Ferric Uptake Regulator (Fur) Mutant of *Shewanella oneidensis:* Possible Involvement of Fur in Energy Metabolism, Transcriptional Regulation, and Oxidative Stress. Genetics and Mol. Biology. (68) 881-892. Doi: 10.1128/AEM.68.2.881-892.2002.

Wang, J., Moolji, J., Dufort, A., Staffa, A., Domenech, P., Reed, M.B., Behr, M.A. (2016) Iron acquisition in *Mycobacterium avium* subsp. *paratuberculosis.* J. Bacteriol. 198:857–866. DOI:10.1128/JB.00922-15

Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Yuan, L., Li, X., Du, L., Su, K., Zhang, J., Liu, P., He, Q., Zhang, Z., Peng, D., Shen, L., Qiu, J., Li, Y. (2020) RcsAB and Fur Coregulate the iron-acquisition system via entC in *Klebsiella pneumoniae* NTUH-K2044 in Response to iron availability. Front. Cellular and Infec. Microbiology. (10) 282. Doi:10.3389/fcimb.2020.00282

Zhou, Q., Su, X., Jing, G., Chen, S., Ning, K. (2018). RNA-QC chain: comprehensive and fast quality control for RNA-seq data. BMC Genomics. (19):144. DOI: 10.1186/s12864-018-4503-6.

CHAPTER SIX: CONCLUSIONS AND FUTURE DIRECTIONS

Conventional wisdom on cultivation *Mycobacterium avium* subsp. *paratuberculosis* (MAP) from primary clinical specimens (such as feces, rectal pinch, environmental samples or milk) suggests that MAP has special iron requirements. For optimal *in vitro* growth, it requires supplementation of the siderophore mycobactin J. Following whole-genome sequencing of MAP K10, multiple computational and functional studies investigated the role of iron in MAP survival (Stratmann *et al.,* 2004; Li *et al.,* 2005; Semret *et al.,* 2005; Alexander *et al.,* 2009; Janagama *et al.,* 2010; Wang *et al.,* 2014). Intracellular concentration of iron must be tightly regulated in order to maintain bacterial cell viability. In bacteria, iron metabolism is under control of two major metallogulators proteins: diphtheria toxin (DtxR) and ferric uptake regulator (Fur). Functional characterization of the iron dependent repressor protein (IdeR) in MAP, a member of DtxR family, showed that it not only controls mycobactin (*mbt*) synthesis but also regulates transcription of genes involved in iron acquisition (*mbtB*), iron storage (*bfrA*) and oxidative stress (Janagama *et al.,* 2009). Furthermore, IdeR has been shown to be significantly different between cattle and sheep strains of MAP (Janagama *et al.,* 2010).

In addition to IdeR, computational analysis of MAP K10 genome, showed the presence of an alternative iron uptake mechanism in conjunction to the presence of two putative Fur boxes (Stratmann *et al.,* 2004). The alternate iron regulator, Fur, is located on a MAP-specific genomic island (LSP15) and has been confirmed to be involved in metal homeostasis. The predicted ferric uptake regulator (*MAP3773c*) has been shown to be downregulated in infected intestinal tissues but upregulated in experimentally infected macrophages (Janagama *et al.,* 2010). Thus, we hypothesized that *MAP3773c* (Fur) encodes a regulatory protein that orchestrates metabolic

pathways required for optimal environmental sensing of iron, providing an alternative iron pathway.

We first performed *in vitro* studies using a transposon mutant of *MAP3776c,* located in the same genomic island (LSP15) as *MAP3773c*. Our data showed that knockdown of *MAP3776c* resulted in the knockdown of *MAP3773c* transcription. Interestingly, complementation of *MAP3776c* restored the transcription of *MAP3776c* and partial cell invasion phenotype. *In vitro* invasion assays of epithelial cell (MAC-T) and its co-culture with monocyte-derived macrophages (MDMs), revealed that the transposon mutant was invasion deficient and was depleted within 48 hr after cell entry. Whether this phenotype was related to lack of *MAP3773c* transcription or a mere response to *MAP3776c* disruption needs confirmation. *MAP3773c* was upregulated in wildtype MAP strains infecting epithelial cells under two different conditions (alone or in coculture with MDMs). Taken together, data presented in Chapter 3 suggests a likely role for LSP15 (and Fur) in *in vivo* and *ex vivo* cell entry and survival suggesting a function in invasion and virulence of this organism.

Next, *MAP3773c,* the Fur ortholog was characterized. We showed a genomic view of the Fur regulatory network, using ChIP-seq analysis. Fifty-eight putative binding sites of Fur were identified both under iron replete and deplete conditions. Three binding regions were identified in genes that are directly related to iron regulation: *MAP3638c* (hemophore-like protein), *MAP3736c* (Fur box) and *MAP3776c* (ABC transporter*).* Using EMSA, specificity and dependence on Mn^{2+} availability for physical binding of Fur to the 19bp consensus Fur box was confirmed. To date, this is the first full characterization of Fur_{MAP}. Previously, other Fur family members in MAP have been investigated, FurA and FurB, also known as Per (peroxidase stress

response) and Zur (zinc uptake repressor) respectively, however no information about their potential roles in iron regulation have been described *(*Eckelt *et al.*, 2014; Eckelt *et al.*, 2015). ChIP-seq analysis of Fur_{MAP} provided potential targets of Fur-mediated gene regulation under iron-stress, however a functional investigation was needed to provide a better understanding on the role of Fur. Thus, an in-frame deletion of Fur was created using homologous recombination. A transcriptional profile of the parent MAP K10, deletion mutant and the complemented strains was developed under iron replete and depleted conditions. Comparison of gene expression differences of mutant against parent and complemented strains showed no differences or patterns that were associated with iron metabolism. Thus, under the current experimental conditions, we are unable to conclude if the lack of transcriptional responses in our study was indicative of a lack of FUR activity.

While transcriptional differences were not informative, the ChIP-seq data combined with the multiple networks linked to iron metabolism still suggest Fur_{MAP} likely plays a role in stress responses. Fur has been extensively investigated in other bacteria, and in addition to its roles with iron-related genes, Fur is also known to regulate genes involved in DNA synthesis, energy metabolism, redox-stress resistance, tricarboxylic acid TCA cycle, and chemotaxis (Escolar, Martin and Lorenzo, 1999; Hantke, 2001; Fillat, 2014). Thus, disruption of Fur regulation is not exclusively related to iron homeostasis – as described in *Klebsiella pneumoniae*, it is possible that our identified Fur_{MAP} binding regions are related to the transcriptional regulation from other cell stresses regulated by Fur_{MAP} rather than being related to iron metabolism (Yuan *et al.,* 2020). Additionally, MAP genome has two iron regulators identified: IdeR and Fur. IdeR was previously proposed to be the main regulator of MAP iron metabolism, so it is possible that IdeR

still regulates most iron-related functions while Fur may be involved in post-transcriptional control or may be triggered by alternative environmental cues (Janagama *et al.,* 2009).

In this thesis, we applied ChIP-seq and RNA-seq, to identify genes regulated by the transcription factor Fur in *Mycobacterium avium* subsp. *paratuberculosis*. In combination with transposon mutant *MAP3776c* and the in-frame deletion mutant *MAP3773c*, we analyzed the relationship between genotype and phenotype of MAP in response to iron stress. While ChIP-Seq and cell invasion data confirmed that LSP15 genes were likely associated with iron regulation in MAP, RNA-seq failed to show any associations. We propose that future studies include a dose-response and a time-course analysis to fully decipher the function of *MAP3773c* and LSP15 genes on iron metabolism in MAP. Titration of iron concentrations in culture media with or without oxidative and acidic stressors will likely provide a clear picture of iron metabolism and potentially the role of FUR in MAP survival.

In addition, investigations targeting functional analysis induced by other stress conditions and understanding the relationship between Fur and IdeR will further our understanding a key virulence determinant of MAP.

REFERENCES

REFERENCES

Alexander, D.C., Turenne, C.Y., Behr, M.A. (2009) Insertion and deletion events that define the pathogen Mycobacterium avium subsp. paratuberculosis. J Bacteriol. 191(3):1018- 25. DOI: 10.1128/JB.01340-08

Brinkworth, A.J., Wildung, M.R., Carabeo, R.A. (2018) Genomewide Transcriptional Responses of Iron-Starved Chlamydia trachomatis Reveal Prioritization of Metabolic Precursor Synthesis over Protein Translation. MSystems. 3(1):184-17. Doi:10.1128/mSystems.00184-17

Eckelt, E., Jarek, M., Fromke, C., Meens, J., Goethe, R. (2014) Identification of a lineage specific zinc responsive genomic island in *Mycobacterium avium* subspecies *paratuberculosis.* BMC Genomics. 15:1076. doi: 10.1186/1471-2164-15-1076

Eckelt, E., Meißner, T., Meens, J., Laarmann, K., Nerlich, A., Jarek, M., Siegfried Weiss, S., Gerlach, G-F., Goethe, R. (2015). FurA contributes to the oxidative stress response regulation of *Mycobacterium avium* ssp. *paratuberculosis.* Frontiers in Microbiology. DOI:10.3389/fmicb.2015.00016.

Janagama, H.K., Senthilkumar, T.M., Bannantine, J.P., Rodriguez, G.M., Smith, I., Paustian, M.L., McGarvey, J.A., Sreevatsan, S. (2009) Identification and functional characterization of the iron- dependent regulator (IdeR) of *Mycobacterium avium* subsp. *paratuberculosis*. Microbiology. 155, 3683-90. doi: 10.1099/mic.0.031948-0.

Janagama, H.K., Lamont, E.A., George, S., Bannantine, J.P., Xu, W.W., Tu, Z.J., Wells, S.J., Schefers, J., Sreevatsan, S. (2010) Primary transcriptomes of *Mycobacterium avium* subsp. *paratuberculosis* reveal proprietary pathways in tissue and macrophages. BMC Genomics. 11, 561. doi: 10.1186/1471-2164-11-561.

Lamont, E.A., Xu, W.W., Sreevatsan, S. (2013) Host-*Mycobacterium avium* subsp. *paratuberculosis* interactome reveals a novel iron assimilation mechanism linked to nitric oxide stress during early infection. BMC Genomics. 2164 DOI:10.1186/1471-2164-14-694

Pandey, M., Talwar, S., Bose, S., Pandey, A.K. (2018). Iron homeostasis in Mycobacterium tuberculosis is essential for persistence. Scientific Reports. Doi: 10.1128/mSystems.00184-17

Semret, M., Alexander, D.C., Turenne, C.Y., de Haas, P., Overduin, P., van Soolingen, D., Cousin, Behr, M.A. (2005). Genomic Polymorphisms for *Mycobacterium avium* subsp. *paratuberculosis* Diagnostics. Journal of Clinical Microbiology. 43, 3704,12. DOI:10.1128/JCM.43.8.3704–3712.2005

Stratmann, J., Strommenger, B., Goethe, R., Dohmann, K., Gerlach, G.F., Stevenson, K., Li, L.,

Zhang, Q., Kapur, V., Bull, T. (2004) A 38-kilobase pathogenicity island specific for *Mycobacterium avium* subsp. *paratuberculosis* encodes cell surface proteins expressed in the host. Infect. Immun. 72, 1265–1274. DOI: 10.1128/IAI.72.3.1265-1274.2004

Thomas, J., Hiltenbrand, T., Bowman, M.J., Kim, H.R., Winn, M.E., Mukherjee, A. (2018). Time-course RNA-seq analysis provides an improved understanding of gene regulation during the formation of nodule-like structures in rice. Plant Mol. Biol. 103:113-128. Doi:10.1007/s11103-020-00978-0

Wang, J., Pritchard, J.R., Kreitmann, L., Montpetit, A., Behr, M.A. (2014) Disruption of *Mycobacterium avium* subsp. *paratuberculosis*-specific gene impairs in vivo fitness. BMC Genomics. 15:415. DOI:10.1186/1471-2164-15-415