# The role of *Mycobacterium tuberculosis* transcription factor Rv1985c in regulating adaptation to hypoxia and reaeration

Marshall Johnathan Davis

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2018

Committee:

David Sherman

Christoph Grundner

Kevin Hybiske

Program Authorized to Offer Degree:

Global Health

University of Washington

**Abstract**

The role of *Mycobacterium tuberculosis* transcription factor Rv1985c in regulating adaptation to hypoxia and reaeration

Marshall Johnathan Davis

Chair of the Supervisory Committee:

David Sherman

Pathobiology Graduate Program

*Mycobacterium tuberculosis* is the deadliest infectious organism in the world, yet remains severely in need of updated treatment options, vaccine candidates, and improved diagnostic tools. This need derives from an incredibly complex host-pathogen interaction that has proved reticent to the identification of clear links between bacterial adaptation and host determinants of infection control. To better understand these links, we have developed a systematic method to evaluate the regulatory role of each transcription factor (TF) in *Mycobacterium tuberculosis (MTB)* during infection. We created a TF-induction library by cloning 207 of the estimated 214 MTB TFs into an anhydrotetracycline-inducible expression vector and transforming these into wildtype H37Rv. This library of mutants was screened through both macrophage infection and a hypoxia/reaeration time course – two critical contexts for the progression of infection. We identified 24 TFs that lead to defective growth in macrophages, as well as 5 TFs with a hypoxia/reaeration induction defect. One TF, Rv1985c, was investigated in further detail and revealed to drive a continued repression of a ribosomal protein locus during the accumulation of the growth defect over the course of reaeration. The regulons identified here could provide novel insight into the specific, evolved transcriptional adaptations required for MTB to cause disease.

**TABLE OF CONTENTS**

**List of Figures**

## List of Tables

**Acknowledgements**

Each and every member of the Sherman lab deserves recognition for their contributions both to this project and my growth and development in graduate school, especially Jessica Farrow-Johnson, Jessica Winkler, Shuyi Ma, and Tige Rustad. Thanks to my committee members for their scientific expertise, flexibility, and support over the course of my graduate career. Heartfelt thanks to my family, who have long been an unbreakable support system and source of inspiration. Finally, I am incredibly grateful to my mentor David Sherman, not only for his help in developing as a graduate student and scientist over the course of this project, but for serving as a wonderful mentor as I embarked upon a career.

## Introduction

The single deadliest infectious organism in the world, *Mycobacterium tuberculosis (MTB)* emerged as a pathogen roughly 70,000 years ago and has had an inestimable impact on humanity ever since[1]. In 2016, MTB killed 1.8 million people and manifested 10.4 million new cases of disease including nearly 500,000 multi-drug resistant cases[2]. In the last 100 years, the total death toll attributed to tuberculosis disease exceeds 100 million[3]. It ranks as the 9th leading current cause of death worldwide, surrounded by better known threats such as diabetes, Alzheimer's, and automobile accidents[2]. Despite its status as a leading global killer, ancient origins, and storied history within our collective cultural conscience, the public health and research outlook surrounding MTB remains one of dire need.

The tools with which we fight tuberculosis include a vaccine from 1921 that neither prevents infection nor adult disease[4], a standard drug cocktail whose components were all discovered prior to 1963, and frontline diagnostics that cannot distinguish infection from active disease. This inadequate set of interventions stems partly from an inability to identify MTB's deterministic virulence adaptations and the specifics of an effective host immune response. Without advances in understanding the molecular etiology of tuberculosis disease it remains unlikely that significant advancements in intervention will occur.

**Mycobacterium tuberculosis.** Tuberculosis is an ancient disease, predominantly thought to have emerged in human populations roughly 70,000 years ago[5-6]. A recently discovered 500,000-year-old *Homo erectus* fossil in Turkey perhaps displays typical MTB bone lesions, potentially indicating an even earlier origin[7]. Regardless of the exact date, tuberculosis emerged and evolved with the human population, and accrued a corresponding cultural impact. Throughout history tuberculosis earned a myriad of descriptive titles, each seeking to appropriately capture the impact of the disease:

consumption, phthisis, white plague, King's evil, scrofula, lupus vulgaris, and Pott's disease all refer to clinical manifestations of the same bacterium. The number of famous individuals known to have suffered from or succumbed to the disease commands perhaps more attention: English kings Edward VI, Henry VII, and Henry VIII, poets John Keats and Elizabeth Browning, composer Frederic Chopin, writers Charlotte, Emily, and Anne Bronte, Fyodor Dostoyevsky, and politicians Eleanor Roosevelt and Nelson Mandela[8]. Famously, in 1882 Robert Koch identified *Mycobacterium tuberculosis* in his experiments to confirm the germ theory via his enduring "Koch's Postulates."

Upon review of our substantial cultural ties to the disease, it seems appropriate that the field discovered MTB to be a pathogen of primary human origin[1]. The MTB complex is composed of pathogens of human (*MTB, M. africanum, M. canettii*), rodent (*M. microti*), seals and sea lions (*M. pinnipedi*), sheep and goats (*M. caprae*), and cattle (*M. bovis*). Recent genetic analyses revealed the most basal ancestor *M. canettii* and the other ancient tuberculosis lineages lack specific genetic deletions carried by the animal lineages, suggesting tuberculosis originally evolved as a human pathogen.

MTB is a gram-positive bacterium belonging to the *Actinomycetales*, though it possesses gram-negative features in its cell wall including peptidoglycan, mesodiaminopimelic acid, and an outer lipid barrier. A feature of Mycobacteria are the mycolic acids: alpha-branched beta-hydroxy fatty acids composed of 70-90 carbon atoms that provide a hydrophobic, waxy exterior[9]. Mycobacteria characteristically possess 60-70% G-C content in their genomes and can be divided into fast and slow-growing species. While *Mycobacterium tuberculosis* take roughly 16-18 hours to double under ideal conditions, the fast-growing soil species *Mycobacterium smegmatis* can do so in only two hours.

**Disease.** Clinically, MTB has proven capable of infecting nearly every tissue within the body, though the vast majority of infections occur in the lower respiratory tract. The classical progression of infection and disease is as follows[4, 10,11]: A diseased individual expels infectious droplets via coughing, sneezing, or speaking, which can contain as few as a single bacterium, and remain suspended in the air for a long period of time until inhalation. The infectious droplets translocate to the lower respiratory tract, where they are phagocytosed by alveolar macrophages. Infected alveolar macrophages then migrate across the lung epithelium to the interstitium. This initial macrophage infection represents the first of a long series of infection stresses to which the bacteria must adapt. While some bacteria are killed by alveolar macrophages, others multiply during this initial phase of infection. Infected alveolar macrophages produce a localized inflammatory response via the release of bacterial cell debris and cytokines, and this will lead to the recruitment of the innate immune response including dendritic cells and monocytes. In mice, over the course of the first 21 days of non-specific response to MTB infection, infected dendritic cells or monocytes traffic to the lymph nodes for adaptive immune response priming. Additionally, the localized inflammatory response in the lung continues to recruit inflammatory monocytes. By the onset of adaptive immunity, the site of infection will be heavily populated by macrophages, neutrophils, and monocytes. Macrophage populations will also differentiate into various classes including multinucleated giant cells and foamy macrophages. Upon the arrival of primed T lymphocytes, infection progresses to the hallmark of TB infection, the granuloma. The specific priming of macrophages by Th1-skewed T cells allows for the control of bacterial replication. The structure of the granuloma classically contains densely packed macrophages and monocytes at the center, surrounded by an influx of lymphocytes on the outer edge. The area will be well infiltrated by capillaries, and a fibrous cuff may also mark the edges of the response. The center of the granuloma often contains a necrotic caseum: an acidic, extracellular environment composed of dead cellular debris that is surrounded by macrophages. The center of the granuloma is characterized by hypoxic conditions that can promote a state of non-

replication and persistence in the bacteria[12]. The majority of infections typically are contained at this stage, only between 5-15% of infections will progress beyond this point[13].

Those granulomas that do progress to active disease are characterized by a liquefaction of the caseous material via hydrolytic enzymes, which promotes a highly replicative environment. Replication continues to the point of granuloma rupture, spilling huge numbers of bacteria into the airway, rendering the individual infectious. Destruction of bronchial structures can lead to the furthering of disease symptoms and dissemination of infection to other sites within the body. This breaking of containment and progression to active disease can occur over a large timeframe – from mere months to decades. The most common risk factors for this progression are HIV infection and/or a compromised immune response. Recent histological and imaging projects revealed the progression of disease varies locally between multiple granulomas within a single individual[14].

**Latent tuberculosis.** As stated previously, most MTB infections are controlled by granuloma formation. Globally, close to two billion people test positive via the tuberculin skin test or interferon gamma release assay with no clinical presentation of symptoms[2] and are considered latently infected with tuberculosis. The greatest risk factor for reactivation of infection and progression to active disease is HIV progression[15]. Significant research efforts have been undertaken to understand what the determinants of granuloma formation and control and potential reactivation are. Increased levels of tumor necrosis factor and interferon gamma are important to the initial establishment of the granuloma, with tumor necrosis factor requiring expression within a specific range to avoid too much or too little inflammatory response[16-18]. Other recent studies have identified latently infected individuals to have higher circulating levels of interferon gamma, lower levels of T regulatory cells, and higher levels of IL-17, IL-23, and Th17 cells, which negatively regulate Tregs[19-21]. Furthermore, a set of studies recently

identified host blood RNA signatures that may predict progression to active disease and treatment outcome[83,84,85]. These signatures showed elevation of type I/II interferon signaling and the complement cascade up to 18 months before disease, which preceded alterations in immune cell subset gene expression changes closer to the outset of disease. Nevertheless, further research into the specific determinants of control and reactivation are required for both the bacterial and immune perspectives.

Clinically, latent tuberculosis infection represents an enormous challenge to global health. The gold standard for infection control, the development of a preventative vaccine, would not impact the current pool of nearly two billion latently infected people. This pool of potential active disease represents a challenge to control efforts, particularly in endemic areas. The identification and treatment of such a broad pool of individuals constitutes a substantial logistical and monetary challenge given current global resources for fighting TB. Better diagnostics, an understanding of disease progression determinants, and new drugs/cocktails all constitute areas of need in regard to latent infection.

**Treatment.** Until the mid-20$^{th}$ century, patients infected with tuberculosis were afforded little recourse against the disease. Prior to the advent of modern medicine, more archaic treatments included blood letting and the laying of hands by a king. Perhaps the most commonly known tuberculosis treatment was the sanitarium. The idea that good air, rest, and proper nutrition could help a patient overcome the disease was widely accepted in the early 1900s, and sanitaria were opened across the world. Actual benefits of this regimen primarily occurred in public health, as this movement effectively quarantined patients with active disease.

The first modern treatment for tuberculosis came in the form of surgical intervention. Procedures focused on collapsing the lung around the infection to deny access to oxygen, and this was

accomplished by the injection of air, oil, or plastic spheres into the pleural space of the lung. As surgical techniques advanced procedures moved towards the removal of ribs or sections of the lungs themselves. More recently, as extensively drug resistant infections increase in frequency doctors have returned to surgical intervention as a last resort[22].

The discovery of the first anti-tubercular agent in 1944, streptomycin, represented an enormous advance for treatment of the disease. Relatively quickly, a host of drugs were identified for treatment of the disease: para-aminosalicylic acid (1949), isoniazid (1952), pyrazinamide (1954), cycloserine (1955), ethambutol (1962), and finally rifampicin (1963). The current treatment regimen consists of rifampicin, isoniazid, pyrazinamide, and ethambutol for two months, followed by an additional four months of isoniazid and rifampicin. This suite of drugs brings with it a burden of side effects including hepatitis, nausea, neuropathy, and renal failure[10]. The length of treatment, host of detrimental side effects, and cost all represent significant challenges in administration of this intervention to the poorer regions of the world where tuberculosis remains endemic today.

The very first clinical trial of streptomycin revealed drug resistance, and the challenge has only continued to grow in the intervening years[23]. Multi-drug resistant tuberculosis (MDR-TB) is defined as a strain resistant to both isoniazid and rifampicin. Globally, there are currently 500,000 MDR-TB cases, and treatment efficacy in these cases falls to only 50%[2]. Extensively drug resistant tuberculosis (XDR-TB) is defined as resistance to isoniazid and rifampicin as well as to any of the fluoroquinolones and one of the other second-line drugs. Cases of XDR-TB have now been observed in India, China, South Africa, Russia, and other eastern European countries[24]. Only one new drug (Bedaquiline) has been approved by the FDA for the treatment of tuberculosis in the past 40 years[25], underscoring the dire need for the discovery of new drugs and identification of novel treatment regimens.

**Vaccine.** The intervention with the potential to most dramatically alter the global landscape of tuberculosis burden is undoubtedly a vaccine. Currently, the approved vaccine for MTB is bacillus Calmette-Guérin (BCG), a strain derived from *M. bovis* and serially passaged in the early 20th century until it became non-pathogenic. BCG is widely administered to infants in areas where tuberculosis disease remains endemic, though its use has been curtailed in areas with limited burden. Among adults, BCG has an efficacy of 0-80%, and is generally thought to be ineffective at preventing pulmonary tuberculosis disease[26,27]. Children less than five years of age, however, show more reliable protection against non-pulmonary disease[28].

While BCG affords some protection to young children, it offers little in terms of reducing the global burden of tuberculosis, and there exists a clear need to develop a better vaccine. This process faces significant challenges, as not only does the field lack a clear picture of the appropriate immune response to elicit with a novel vaccine, but there remain questions surrounding the limited mechanism of protection in BCG[29]. Nevertheless, in 2016 13 vaccine candidates underwent phase I or later clinical trials[10]. Moving forward, critical questions relevant to the design of these vaccines must be addressed. These include defining the specific contribution of CD4 and CD8 T cells to immunity, the role of innate lymphoid cells (including mucosal associated invariant T and invariant natural killer cells), the most appropriate immune response to achieve via adjuvant, and whether post-exposure vaccination could function as a form of treatment[29].

**Chapter 1. TRIP screen for regulons with a macrophage phenotype**

**Introduction**

The macrophage acts as the first intimate contact of the immune system with MTB. Following inhalation and deposition in the lung, the bacteria are first phagocytosed by alveolar macrophages. Infected alveolar macrophages migrate through the lung epithelium to the interstitium, where a local inflammatory response builds. Over time, an influx of lung interstitial macrophages occurs and even as the adaptive response shapes granuloma formation the macrophage remains the primary contact point for direct immune system interaction with the bacteria. Given this prolonged, stress-ridden interaction, MTB has long been forced to evolve specific, coordinated responses to survive the macrophage environment.

Throughout infection, MTB works to subvert and modulate the immune response to its benefit. Under normal circumstances, a phagocytosed bacterium is trafficked to an acidic, hydrolytic environment within the lysosome. MTB has evolved the ability to subvert phagosome maturation, manipulating the endosomal compartment to an only slightly acidic pH, low influx of lysosomal hydrolases, and maintained connection to the endosomal pathway[30-32]. The bacteria also disrupt the phagosomal membrane via the ESX-1 secretion system, leading to secretion of bacterial products and DNA into the cytosol – though there is contention MTB may fully escape the cytosol in some cases. The predominant hypothesis as to the function of signaling into the cytosol of the macrophage is the intentional activation of the type 1 interferon response, a classical mediator of viral response that conversely leads to worse outcome in bacterial infections[33]. Mice lacking type 1 interferon receptor carry significantly lower bacterial burdens, and the mediators of this outcome are thought to be decreases in key MTB-response cytokines (TNF, IL-1B), lowered IFN-γ mediated activation of macrophages, and IL-10 driven immunosuppression[34-36]. Finally, differential immune cell death pathways have been associated with

bacterial burden and may act as a key determinant of disease progression[37-41]. Caspase-mediated cell death including autophagy and apoptosis is associated with bacterial control, while necrotic death increases bacterial burden and disease progression.

Significant efforts have been made to characterize MTB's transcriptional response and essential requirements for the infection of macrophages. Schnappinger and associates produced a seminal study defining the intraphagosomal transcriptome of MTB, providing an overview of the expression changes accompanying macrophage infection[42]. The group defined 454 induced MTB genes and 147 repressed associated with infection at 48 hours, and the set of genes was enriched for functional categories including DNA-damage, cell wall damage, and fatty acid metabolism. Similar studies have extended transcriptomic analysis to clinical isolates and environmental cues in an attempt to define the critical transcriptomic landscape for macrophage infection, and additionally identified genes enriched for hypoxia, oxidative stress, and cell wall and DNA repair[43-45].

Despite these significant advances in understanding the transcriptional landscape during macrophage infection, it remains challenging to link specific transcriptional changes to the host-pathogen interplay outlined earlier. These broad studies provide relatively little information concerning how these transcriptional responses are specifically controlled, nor do they identify how particular effectors are linked in their importance to infection. Defining specific gene subnetworks that determine macrophage infection and disease progression is one focus of the Sherman lab. We propose to better identify and study the specific transcriptional changes of MTB that are critical for disease progression by studying the regulators of the transcriptome – transcription factors. To this end, the lab created a pool of inducible transcription factor expression mutants and previously published DNA-binding networks and microarray expression analysis of each individual induced mutant during log phase growth[46,47]. These mutants can

be grown in a pool and tracked via sequencing a genetic tag unique to each (Figure i). In the absence of induction, these strains grow at similar rates to wild type within the pool in response to stress. Upon induction, the exposure to stress will produce differential growth phenotypes in a subset of mutants, and sequencing will detect changes in their relative abundance to the rest of the pool. This method is referred to as the transcriptional regulon induced phenotype (TRIP) assay.

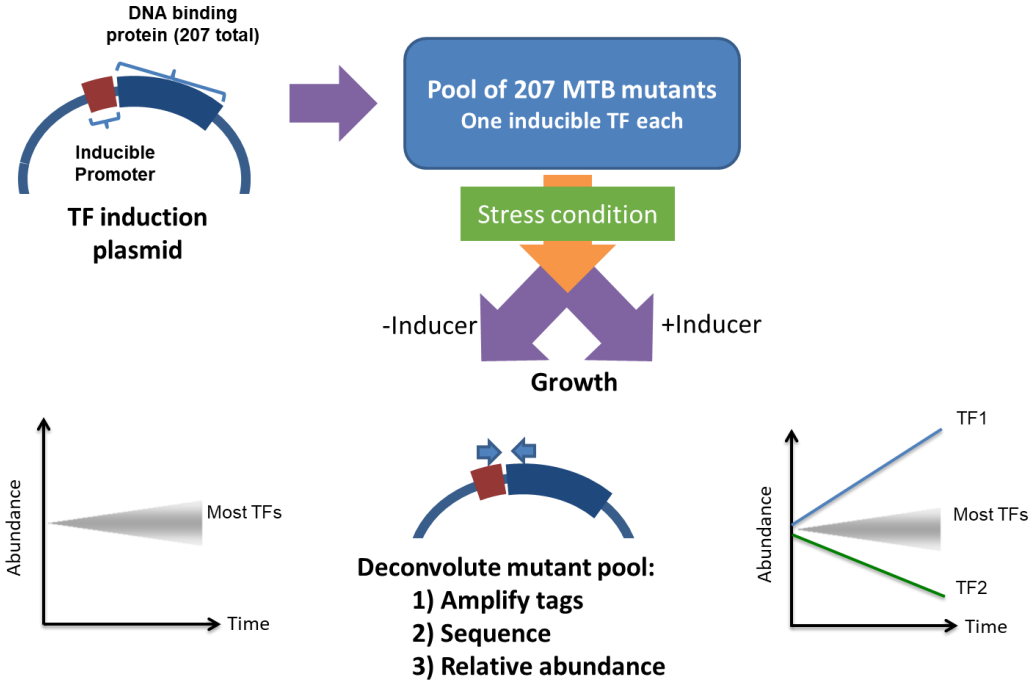



Figure i. Overview of the TRIP method.

The identified phenotypes stem from a single induction event propagating through its specific, evolved set of network interactions to culminate in a differential stress outcome. Identified phenotypes can then be further characterized transcriptionally to define those regulatory interactions and effectors key to adaptation to a given stress. We hypothesize induction events and their sustained interactions over the course of the experiment will be less disruptive over time to the network than a traditional gene knockout, allowing for the identification of regulators and regulons previously unassociated with a given stress by gene disruption methods. Here we undertook the initial application of this TRIP pool to a

stress context in the form of macrophage infection, with the goal of identifying transcription factors and regulons that drive a growth defect when induced throughout macrophage infection.

## Methods

**TRIP pool construction.** Construction of the induction mutant strains has been published in detail in both Minch et al. and Rustad et al[46,47]. In brief, for each of the 214 putative DNA-binding proteins in MTB we searched a Gateway Entry Clone library of open reading frames in the pDONR221 backbone (PFGRC/Colorado State University under NIAID contract HHSN266200400091c, available from BEI). A total of 206 TF entry clones were successfully created and cloned into a vector via a Gateway cloning recombination cassette (gift of Eric Rubin) that placed the TF under the control of a tetracycline inducible promoter. A C-terminal FLAG tag was also added to each construct (except for Rv3133c/DosR, which is N-terminal), before transformation into MTB H37Rv. The final ATc-inducible strains are available from the BEI strain repository (NR – 46512). To construct the TRIP pool, all strains were cultured until stationary phase (OD 1.0) and diluted back to OD 0.1 in a cultured containing each of the 206 strains. Cultures were stored in 1 ml aliquots at 50% glycerol concentration at -80 $^{o}$C.

**Standard culturing conditions and log-phase experiments.** All strains were cultured in Middlebrook 7H9 with ADC supplement(Difco), 0.05% Tween80, and under constant agitation at 37 $^{o}$C. TRIP strains containing the ATc-inducible expression vector were additionally cultured with 50 μg/ml hygromycin B to prevent plasmid loss. Log-phase experiments were performed under aerated conditions, under constant agitation, and growth was monitored by OD600. 18 hours (approximately one cycle of replication) before the day 0 time point cultures in early log phase (OD 0.1 – 0.3) were split to OD 0.05 and ATc (100 ng/ml final concentration) added to one. Cultures were allowed to continue an additional 5 days before the final time point was collected in the log-phase experiments.

**THP-1 culture and differentiation.** For our macrophage model of infection we utilized human-macrophage-like THP-1 cells. THP-1 cultures were grown in PRMI 1640 (ThermoFischer Scientific), supplemented with 10% fetal bovine serum (FBS), at 37 $^{o}$C, 5% $CO_2$. For quality control, THP-1 cells were not used in experiments past ~10 passages from thaw, to ensure uniform morphology and cell health within the cultures. Undifferentiated THP-1 cell concentration was maintained between 1.0 x $10^5$ and 1.0 x $10^6$ to ensure health. To differentiate cells into the macrophage-like attachment phenotype, cells were centrifuged at 1800 rpm for 10 min and re-suspended at a concentration of 8.0 x $10^5$ cells and 100 nM phorbol 12-myristate 13-acetate (PMA) (Sigma). Treated cells were then dispensed in 5 ml aliquots into vented T25 flasks (Corning) and allowed to rest for 72 hours. Following differentiation, PMA-treated media was removed, cells were washed with fresh RPMI + 10% FBS, and allowed to rest for an additional 24 hours before infection.

**THP-1 infection and TRIP screen.** 18 hours before THP-1 infection, a log phase culture of the TRIP pool was split into two 10 ml cultures at OD 0.1. The induced culture was also brought to a concentration of 100 ng/ml ATc. Each replicate in our macrophage screen consisted of a T25 flask seeded with 4 million differentiated THP-1 cells. For infection, the early log phase TRIP uninduced and induced pools were re-suspended in RPMI + 10% FBS at a concentration of 2.5 x $10^6$ bacteria/ml. 2 ml of TRIP dilution (multiplicity of infection of 1) were then added to each flask and allowed to infect at 37 $^{o}$C for 4-6 hours. Following infection, each flask was emptied of media, washed thoroughly with warm PBS three times to remove any extracellular bacteria, and resupplied with 5 ml RPMI + 10% FBS (+/- 100 ng/ml ATc). Cultures were returned to 37 $^{o}$C, 5% $CO_2$ for the duration of macrophage culture. For each time point, DNA from 3 separate samples was collected immediately following infection and again after five days of intra-macrophage growth.

**DNA isolation.** Macrophages were lysed using 1 ml 0.1% Triton X-100 in PBS for 10 minutes at 37 $^{o}$C. Lysed cells were collected, spun down for 5 minutes at 3200 g, and supernatant removed. The pellet was resuspended in 0.5 ml 1X TE, pH 8.0 and transferred to a lysing matrix B tube (MP Bio). Samples were bead beat 3 times for 30 seconds at 6.5 speed to lyse bacteria using the OMNI Bead Ruptor 24 Elite bead beater. Lysed samples were then spun down for 1 minute at 13,000 x g and 300 ul supernatant transferred to a clean o-ring screw cap tube. Samples were then sterilized for 30 minutes at 95 $^{o}$C before removal from the biosafety level 3 facility. DNA extraction was completed using the standard kit protocol for the MagJET Genomic DNA Isolation kit (ThermoFisher). Final DNA yield was assessed by Nanodrop (ThermoFisher).

**Illumina library prep and sequencing.** To amplify the unique genetic tags of each mutant for Illumina sequencing, we performed a limited PCR amplification using a pool of 207 primer pairs corresponding to each specific TRIP mutant tag. 5 ul of mixed macrophage and MTB DNA were combined with 5 ul of primer mix (37.4 uM), and 10 ul 2X PCR Master Mix (ThermoFisher) and amplified under the following cycling conditions for only 20 total cycles: 95 $^{o}$C for 10 min, 95 $^{o}$C for 15 s, 60 $^{o}$C for 4 min, repeat step 2 19x. Following this, the ~220 bp unique genetic sequences were isolated using size selection via Agencourt AMPure XP beads (Beckman Coulter). Samples were prepared for Illumina sequencing following the standard library prep protocol in the NEBNext Ultra DNA Library Prep Kit for Illumina. Final library concentration was quantified via KappaQuant (Kapa Biosystems). Sequencing was completed at the Northwest Genomics Center at the University of Washington using a NextSeq 500 platform. Returned short sequencing read FASTQ files were aligned to the H37Rv reference genome and analyzed for read counts per gene. Each gene was normalized for the total number of reads per sample (each lane of the sequencing run will contain a differential number of total reads), as well as individual gene

length (larger genes will obtain more reads). Data processing was completed in R, using the edgeR and limma packages.

## Results

**Establishment of log phase induction phenotypes.** Before investigating regulons driving a phenotype in a given stress, we sought to identify those transcription factors that possess a phenotype due to induction during log phase growth. The TRIP pool was grown to OD 0.1 under standard culturing conditions with continuous agitation. The induced pool was then treated with 100 ng/ml ATc and both the uninduced and induced pools cultured an additional five days, at which point DNA isolation and library preparation were carried out (see Methods).

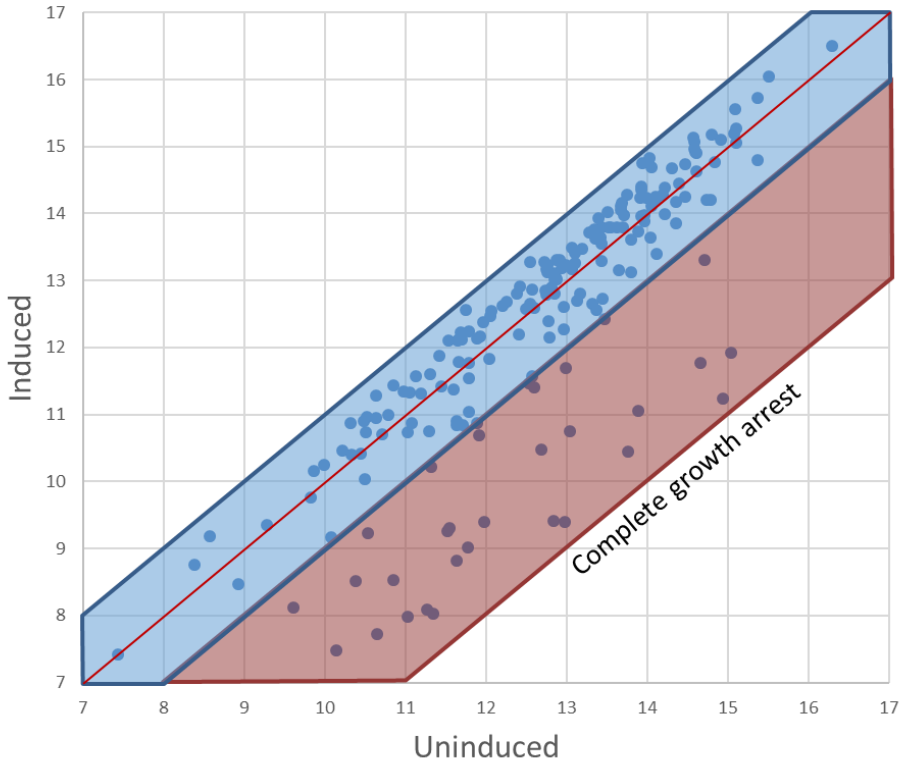



Figure 1. Relative abundance of induced and uninduced TRIP strains after five days of growth. Values on each axis are the log2 number of reads per gene per one million total sample reads. Each data point represents a single TRIP strain. This log phase control data set is composed of multiple independent experiments, and strains are represented by 9 induced and 17 uninduced experiments. The blue zone denotes < 2-fold difference, red ≥ 2-fold difference.

Of the 207 strains within the mutant pool, 24 TRIP strains possess at least a two-fold greater defect in the induced compared to uninduced condition after 5 days of growth (Figure 1). A two-fold difference is equivalent to one less doubling in bacterial growth. Assuming a standard doubling time of roughly 18 hours, this 5-day time course would allow ~6 doublings in the uninduced strains. Those strains with the largest defects have a nearly 6-doublings difference, which is equivalent to complete inhibition of growth within one cycle of replication following ATc addition. TRIP strains falling along the red line and within the blue shaded region in figure 1 possess no difference between the induced and uninduced conditions, and this is the case for 169 of the induction strains during log phase growth.

**Macrophage Defects.** Subsequently, we set out to characterize those TRIP strains that cause a defect over the course of macrophage infection.

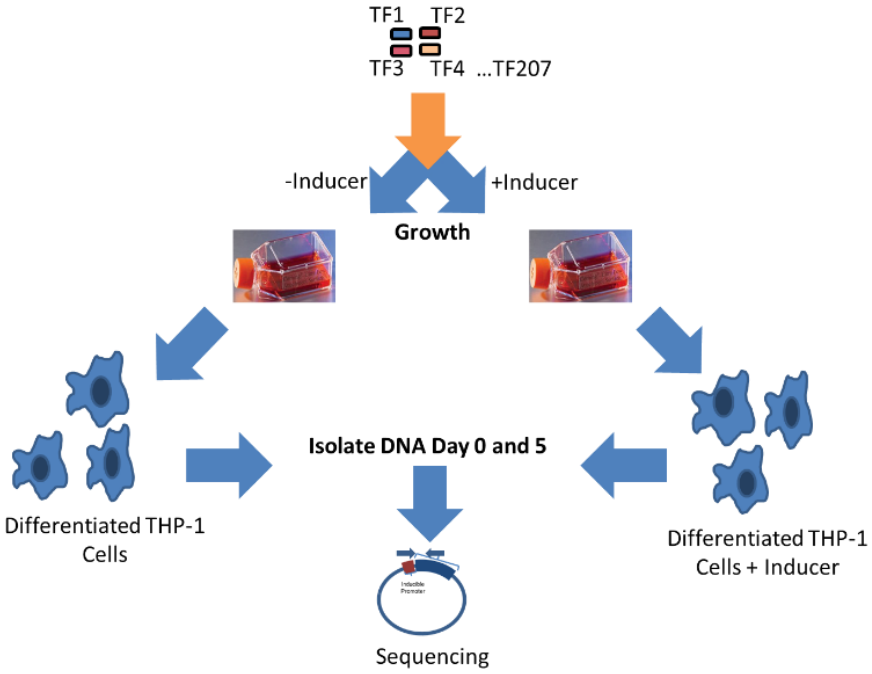



Figure 2. Overview of the macrophage TRIP screen.

Each replicate in the TRIP macrophage screen was a T25 flask seeded with 4 million THP-1 cells and differentiated into the macrophage-model phenotype via treatment with PMA for 72 hours. The TRIP pool was grown to early log phase growth (~OD 0.2), split roughly 18 hours before infection, and one half induced with 100 ng/ml ATc. On the day of infection, the induced and uninduced pools were diluted into RPMI and the flasks infected at an MOI of 1 for 4-6 hours. Following infection, the flasks were washed 3 times with warm PBS before being fed back with 5 ml RPMI (+/- ATc) and incubated at 37 $^{o}$C, 5% $CO_2$ for 5 days. For sample collection, the flasks were treated with 0.1% Triton X-100 for 10 min to lyse the THP-1s cells, at which point the lysate was collected and processed through our standard DNA isolation and library prep procedure (see methods). This screen was successfully completed three separate times, with each run containing three biological replicates per time point.

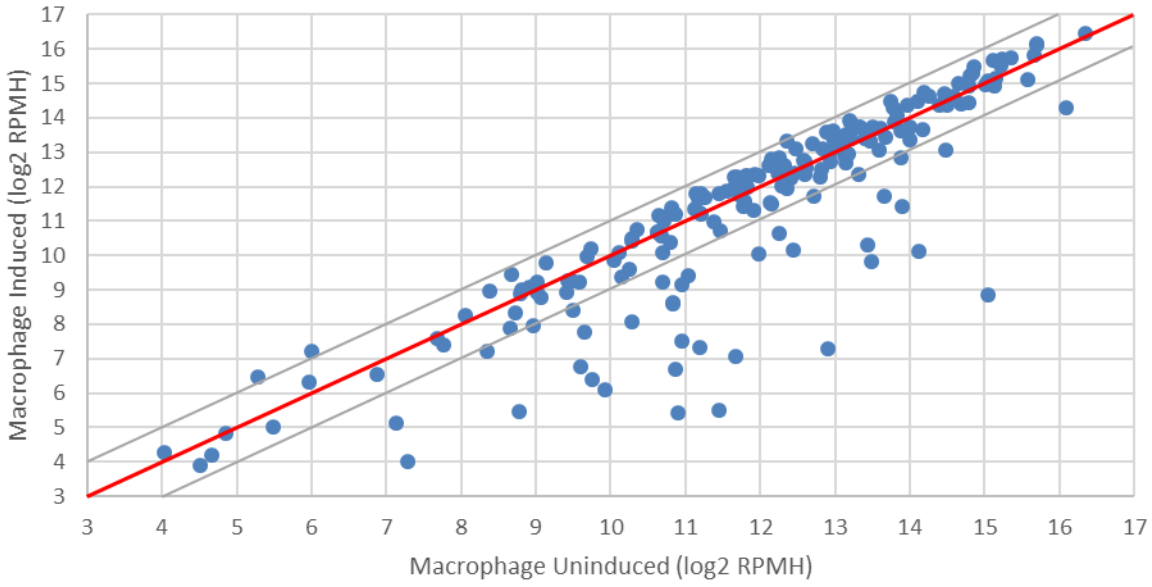


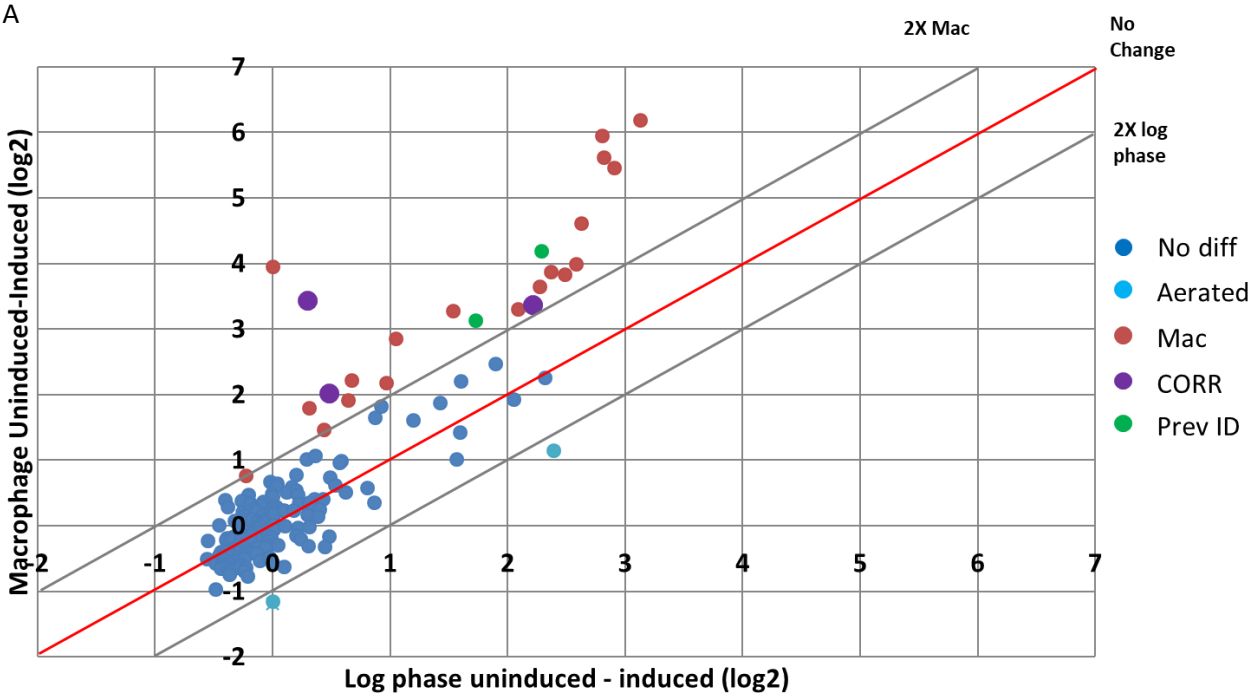

Figure 3. Day 5 induced vs uninduced macrophage phenotype comparison. Graphed values are log2 reads per million hits. Each data point represents a single TRIP induction strain. The data points here are taken from an average of three total screens and comprised of 9 biological replicates in both the induced and uninduced conditions.

Much like the log-phase induced vs uninduced comparison, the majority of TRIP strains do not show a defect over the course of macrophage infection (Figure 3). 168 strains fall along the red line and have a less than 2-fold difference between the induced and uninduced. Two transcription factors, Rv3060c and

Rv1379 fall directly on the 2-fold cutoff for increased growth in the induced condition. 36 strains possess a greater than 2-fold defect in relative abundance after 5 days of growth within THP-1 cells. These strains fall below the two-fold cutoff and are skewed towards the uninduced axis.

**Comparison of macrophage and log-phase phenotypes.** Our primary goal for this screen was to identify regulons that have a defect unique to the macrophage environment, and to identify this population we compared the defects identified in the macrophage to those that occur during log-phase growth. To do this we quantified the defects in each condition by subtracting the induced reads from the uninduced at five days of growth, and then graphed them against one another. In short, the macrophage (Figure 3) and log phase (Figure 1) day 5 growth differences are now graphed against one another. Here, those TRIP strains with an equivalent defect between the two experiments will fall along the central red line, and our phenotypes of interest will fall above a 2-fold cutoff.

A



B

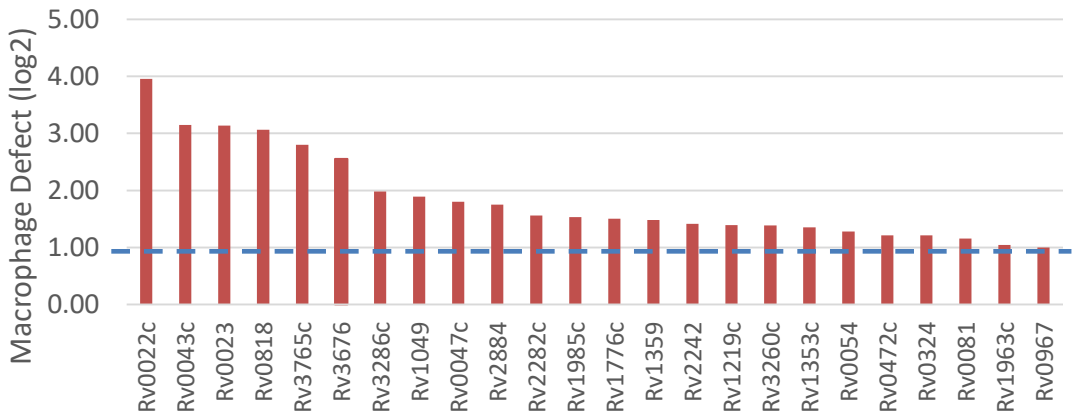



Figure 4a. Comparison of macrophage and log-phase phenotypes to identify regulons driving a defect unique to the macrophage environment. Data points each represent a specific TRIP strain and are a comparison of the uninduced – induced log2 reads per million hits for the macrophage and log-phase data sets. Strains lacking a phenotype are in blue, log-phase increased growth in light blue, macrophage defect regulons are red, CORR regulons purple, and previously identified macrophage regulons in green. b. Quantification of macrophage regulon defects by strain. This graph depicts the magnitude of defect relative to log phase growth for each strain. The blue dashed line represents the two-fold difference significance cutoff.

This comparison succeeded in identifying TRIP regulons with a greater defect resulting from induction within the macrophage environment (Figure 4a). 163 strains did not possess an induction defect in either context. 13 regulons did have a significant difference between the induced and uninduced conditions over 5 days of macrophage growth, but it did not differ between log phase and macrophage infection induction (red line). Three phenotypes were increased during log phase growth relative to macrophage infection (light blue). A total of 24 target regulons were identified (red). These strains possess at least a two-fold greater defect over the course of a five-day macrophage infection relative to the defect accrued during five days of log phase induction (Figure 4b). The enlarged, purple markers signify three strains also identified by our hypoxia/reaeration screen (see chapter 2). The green dots represent Rv1049 and Rv1219c, both TFs have previously published results indicating an association with macrophage outcomes and serve as a validation of our results[48,49]. To further validate our identified phenotypes, we performed a z score comparison between the macrophage and log-phase phenotype data sets. A z score estimates the number of standard deviations a measurement falls from a

comparative population, and in this case was calculated for a given mutant as [(mean macrophage uninduced – induced) – (mean log phase uninduced - induced)] / (standard deviation all log-phase uninduced – induced). Across all mutants, the median number of standard deviations the macrophage phenotype fell from the log phase was 0.55. For our macrophage defect regulons, the median difference was 9.62 standard deviations, and all regulons fell at least 3 standard deviations from the log phase phenotype. Three standard deviations is the equivalent of 99.7% of all data within a given population, and here represents a second significance cutoff for our identification of macrophage growth defects.

**Leveraging our published TRIP data sets.** We turned to the Sherman Lab's previously published data sets in order to better characterize our newly identified macrophage defect regulons[46,47]. The expression data allowed us to identify all transcriptional changes driven by a TRIP regulator, and the ChIP data set identifies all specific binding locations throughout the genome for a given transcription factor.

Examining published functions, the macrophage defect regulons fall into four categories. 13 of the regulators have no published functional significance in macrophage infection, 4 have been previously associated with virulence or adaptation to stress, and two are known to mediate drug resistance or tolerance. We have also identified 4 DNA-binding proteins that do not carry out transcription factor activity but still drive a macrophage defect. The amount of log-phase expression changes mediated by the macrophage TRIP strains varies greatly, and those TRIP regulators lacking in published function also possess the lowest number of log-phase expression changes. The number of ChIP binding peaks for a

given regulon does not correlate with either its functional category or number of log-phase transcriptional changes.

Table 1. Log-phase expression, chromatin immunoprecipitation, and publication results for the macrophage infection regulons.

| Macrophage Regulon | Name | log2 defect | Ind | Rep | ChIP peaks | PubMed |
|---|---|---|---|---|---|---|
| **No Previous Macrophage Association** | | | | | | |
| Rv0043c | | 3.1 | 11 | 6 | 6 | no published results |
| Rv0818 | | 3.1 | 323 | 231 | 21 | no published results |
| Rv3765c | tcrX | 2.8 | 54 | 40 | 85 | two-compoenet regulator expressed during iron stress |
| Rv2884 | | 1.8 | 41 | 134 | n/a | no published results |
| Rv2282c | | 1.6 | 22 | 0 | 1 | no published results |
| Rv1985c | | 1.5 | 1 | 3 | 231 | lysine dependent regulator, no functional characterizatioin |
| Rv1776c | | 1.5 | 12 | 11 | 72 | no published results |
| Rv1359 | | 1.5 | 0 | 0 | 3 | no published results |
| Rv2242 | | 1.4 | 29 | 62 | n/a | no published results |
| Rv1353c | | 1.4 | 44 | 50 | 413 | no published results |
| Rv0472c | | 1.2 | 1 | 15 | 8 | no published results |
| Rv1963c | mce3R | 1.0 | 12 | 2 | n/a | negatively regulates membrane lipid transporters |
| Rv0967 | csoR | 1.0 | 0 | 4 | 331 | copper homeostasis |
| **Virulence/Stress Regulator** | | | | | | |
| Rv0022c | whiB5 | 4.0 | 216 | 219 | 95 | null mutant linked to virulence in both short-term and chronic mouse infection |
| Rv0023 | | 3.1 | 489 | 404 | 241 | CORR. See chapter 2 of thesis |
| Rv1049 | | 1.9 | 4 | 70 | 174 | oxidation sensing regulator shown to be important for macrophage survival |
| Rv0081 | | 1.2 | 283 | 119 | 647 | CORR. See chapter 2 of thesis |
| **Drug Resistance** | | | | | | |
| Rv1219c | | 1.4 | 41 | 29 | 37 | efflux pump regulator |
| Rv0047c | | 1.8 | 29 | 33 | 348 | resistance to putative antitubercular agent |
| **Non-Transcription Factor Function** | | | | | | |
| Rv3676 | | 2.6 | 43 | 42 | 11 | cAMP receptor protein, potentially linked to persistence |
| Rv3286c | sigF | 2.0 | 234 | 71 | 13 | sigma factor involved in stress response, drug target |
| Rv3260c | whiB2 | 1.4 | 104 | 16 | 16 | chaperone-like function, associated with cell division |
| Rv0054 | ssb | 1.3 | 2 | 16 | 5 | single stranded binding protein |

## Discussion

In this experiment we set out to determine whether our transcription factor induction method could discern regulons specific to the adaptation to a given stress. We first needed to identify those regulons that drove a defect due to induction under standard culture conditions, in order to compare with our stress-generated defects. We found 24 TRIP strains possessed a log phase growth defect after 5 days (Figure 1). These strains serve as a control moving forward, as their failure to repeat in any given experiment could indicate a method failure rather than a lack of identified phenotypes. We next

quantified 36 strains with a growth defect within macrophages after 5 days of growth (Figure 3). Comparison revealed 13 of these regulons were equivalent to the defect incurred during log phase growth (Figure 4). However, 24 regulons possess an induction defect after five days of growth in the macrophage that is at least two-fold greater than an equivalent log phase phenotype.

We note the 13 regulators identified that have not previously been linked to adaptation within the macrophage. In the introduction, a comparison was made between this new TRIP screen and more traditional gene disruption or knockout methods. These regulons, 9 of which lack any sort of published characterization whatsoever, potentially represent evolved sets of coregulated genes that respond to macrophage infection that have been missed by previous gene disruption libraries screened through macrophages. While we are aware of no other experiments specifically identifying a macrophage infection phenotype, the published functional characterizations of Rv3765c, Rv1963c, and Rv0967 appear to fit plausibly with the known stress systems surrounding MTB's infection of the macrophage. During macrophage infection, the host works to sequester essential metals, including iron, from the pathogen[50]. Rv3765c is a two-component response regulator known to be upregulated under conditions of iron-limited stress[51]. While Rv3765c's specific downstream regulatory activities related to iron homeostasis are not known, the identification of this regulon represents an opportunity to push forward the field's understanding of adaptation to iron sequestration in the macrophage. Similarly, Rv0967 (*csoR*) has been shown to be a critical regulator of copper homeostasis. At low levels of environmental copper, the loss of *csoR* led to a bacterial stress response characterized by the induction of the *dosR* regulon[52]. Further characterization of this regulon may be warranted in light of recent studies indicating the accumulation of toxic levels of copper in the MTB phagosome is a macrophage defense strategy[53]. Finally, Rv1963c is defined as a repressor of a family of membrane lipid transporters, and its knockout led to alterations in the lipid content of MTB[54]. The importance of lipid uptake and

utilization, particularly as it relates to lipid-loaded foamy macrophages, is an ongoing area of investigation and should be considered relative to this regulon[55,56].

Another functional association of interest identified in our macrophage regulons is the mediation of drug resistance in macrophages. The only known publication for Rv0047c identifies it as containing mutations in resistance mutants to a novel antitubercular compound[57]. While not itself particularly noteworthy, this underscores the point that each of the regulators here represents potentially a novel therapeutic target against MTB and that the elucidation of its macrophage specific regulon could help inform future drug regimen design. More concretely, Rv1219c is known to repress the efflux pumps Rv1217c-1218c, which have been shown to mediate the efflux of a broad number of structurally unrelated classes of antitubercular compounds[48,58]. Understanding how and to what end this regulon is activated could play an important role in designing future chemotherapeutic interventions targeting macrophage infection.

Finally, we identified three strains here we have labelled as CORRs that are highlighted in Figure 4. These strains were also identified in our hypoxia and reaeration experiments in chapter 2 and will be discussed in further detail there.

**Future Work.** A key biological question for further analysis is whether the macrophage context imparts differential regulon expression in those induced transcription factors with a defect. Table 1 highlights a wide diversity in the number of specific DNA-binding events detected by chromatin immunoprecipitation, often in excess of the number of corresponding log phase expression changes. We hypothesize these observed quiescent binding events could mediate expression changes within the correct environmental context, and therefore predict the induction of a given TF may drive a differential

regulon within the macrophage. As such, it is imperative we define the macrophage-specific regulons of these transcription factors in order to better understand how the altered transcriptional subnetworks ultimately manifest in a growth phenotype.

The investigation of the TRIP screen within the context of different macrophage subsets and outcomes could also provide novel associations between bacterial transcriptional networks and infection outcomes. Screening the TRIP system through models of differential cell death pathways, immune activation axes, and macrophage lineages could provide causal links between bacterial adaptation and the few known correlates of disease progression the field has identified.

**Chapter 2. TRIP in hypoxia, reaeration, and an investigation into Rv1985c**

**Introduction**

Chapter 1 of this thesis investigated those regulons critical to macrophage adaptation, as it is the most intimate contact point between MTB and the host. However, a myriad of other factors play a role in the progression to active disease – chief among them the granuloma. Following infection of alveolar macrophages and a period of unchecked replication, the influx of the adaptive immune response drives the formation of granulomas and control of infection. Granulomas are highly stratified structures, in which bacteria are directly surrounded by densely packed macrophages and monocytes, which are further ringed by the influx of lymphocytes. In the majority of cases, the formation of granulomas around infection is sufficient to prevent the transition to active disease, and these individuals are considered to be latently infected. Currently, there are thought to be roughly 1.8 billion people with latent infection, and these individuals compose an enormous pool of potentially active cases[2].

Major questions within the field are the determinants of active disease progression out of the granuloma, however, certain immune correlates have been identified. Sufficient levels of IFN-γ and tumor necrosis factor are key determinants of successful granuloma formation and maintenance[16-18], and the single largest predictor of reactivation disease is HIV infection,[59,60] which increases the chance of active disease progression by 8-10% per year. Other immune factors recently associated with control of disease include higher levels of circulating IFN-γ in latently infected individuals, lower numbers of circulating T regulatory cells, and higher levels of Th17 cells[19-21]. While these broader immune correlates represent steps forward, recent work has highlighted that lesions within the same lung experience diverse disease outcomes, which likely are the result of local phenomena[14]. Finally, recent publications have identified host blood RNA signatures that may help predict treatment outcome and disease progression[83,84,85].

A host of factors are thought to contribute to the control of MTB infection – including nutrient limitation, acid stress, and nitric oxide exposure – but perhaps the most important stress within the granuloma is hypoxia. MTB classically infects the most oxygen-rich sites of the body, and reactivation disease occurs most commonly in the upper lobes of the lungs – the most oxygen rich site of the body[61]. Treatment regimens have also highlighted the importance of oxygen tension. The standard TB treatment before the antibiotic era consisted of surgery to collapse the lung and deprive the bacteria of oxygen, and another study demonstrated that granulomas resistant to chemotherapeutic treatment were associated with direct connection to open airways[62]. Finally, the famous sanitarium movement could potentially be explained on a physiological level by lower oxygen tension leading to decreased disease and transmission at higher altitudes where many facilities were located[63].

Studying the effects of hypoxia within the granuloma itself proves difficult, primarily due to a lack of readily available and physiologically relevant animal models. Commonly used mouse strains do not form hypoxic granulomas and, in contrast to human disease, granulomas typically are not well-formed structures and lack the important traits of caseation and cavitation[64,65]. Guinea pigs and rabbits do form hypoxic granulomas resembling those in humans, but they lack a latent stage[65]. Non-human primates offer strong models of human disease, including hypoxic lesions and latent infection, but due to cost and difficultly have not been widely utilized for these studies[66,67]. Alternatively, *in-vitro* hypoxic culture models have been identified and the MTB transcriptional response to hypoxia has been thoroughly studied within the Sherman lab. The initial response to hypoxia is mediated by the transcription factor DosR, which rapidly induces a set of ~ 50 genes[68]. The activation of DosR is carried out by the histidine kinases DosS and DosT, which utilize heme prosthetic groups to sense redox state and oxygen tension, respectively[69,70]. However, the DosR regulon has increasingly been shown to mediate a more general

stress response, and cannot explain MTB's adaptation to hypoxia. Further characterization of hypoxic time points revealed the induction of the enduring hypoxic response, a set of 230 genes that appear between 4 and 7 days of hypoxic culture[71] independently of the DosR regulon.

These two sets of genes provide the transcriptional landscape of MTB adaptation to hypoxia, but as with the macrophage, further parsing of this network could provide important insight. While there are thirty transcription factors found within the enduring hypoxic response, it has yet to be determined how specific subnetworks of the response are coordinated, whether there are other regulators yet to be identified, and to what degree subsets of genes are redundantly regulated. The identification of a specific subset of regulators coordinating the hypoxic response could provide key insight into the networks of evolved transcriptional adaptation required for the establishment of latency and progression to active disease. Moreover, a core set of hypoxic regulons would also be a valuable tool in investigating outcomes within disparate lesion physiologies or host blood RNA signatures.

Accordingly, we set out to define the core set of regulators and regulons required for surviving hypoxia and reaeration by screening our TRIP strains in a hypoxic/reaeration culture model. In addition to the identification of key regulons, here we investigate the function of Rv1985c as revealed through mutant phenotypes and other analyses.

**Method**

For a detailed discussion of TRIP strain construction, culturing conditions, and DNA isolation protocol see the methods section for chapter 1.

**Hypoxic culture.** Hypoxic culture conditions were achieved using a 1 L three-armed spinner flask (Corning), pictured in Figure 5. The flask featured a magnetic stir paddle and two spouts, all of which were sealed using o-ringed caps to prevent unwanted gas exchange. One spout was reserved for sample collection, and the other featured a cap with two open metal tubes for gas exchange. Both tubes were capped with 0.2 micron filters for biocontainment and to prevent contamination. Hypoxia was induced by flowing low oxygen gas (0.2% $O_2$ with $N_2$ balance) at a rate of 0.15 sq ft/min. Cultures were constantly agitated at 60 rpm for the duration of hypoxic culture inside of a 37 $^{o}$C incubator.

**Hypoxia/reaeration TRIP screen.** Early log phase cultures of the TRIP pool were split and one half induced with a final ATc concentration of 100 ng/ml 18 h before hypoxia was induced. Cultures were diluted to OD 0.1 and 400 ml aliquots transferred to hypoxia flasks. Day 0 time points were collected immediately preceding low oxygen gas flow and the induction of hypoxia. Cultures were kept under constant hypoxic conditions for 7 days, and only removed from gas flow/incubation briefly if sampling was required. For reaeration, samples were removed from the hypoxia flasks and returned to standard culturing conditions as described in chapter 1 for an additional 5 days. For the TRIP screen, samples were collected and genomic DNA isolated at day 0 before hypoxic induction, following 7 days hypoxic culture, and after 5 days reaerated growth. Individual follow up experiments collected samples for CFU enumeration and RNA isolation at day 0 before hypoxia, days 2, 4, and 7 of hypoxia, and each of the five days of reaeration.

**RNA isolation.** Samples were pelleted via centrifugation at 3700 rpm for 5 min, 4 $^{o}$C. Pellets were immediately resuspended in Trizol, transferred to a lysing matrix B tube (MP Bio), and bead beat 3 times for 30 seconds at 6.5 speed, resting on ice for a minimum of 30 seconds in between runs (Omni Bead Ruptor 24 Elite). The mechanically lysed mixture was then centrifuged at max speed for 1 min and the supernatant transferred to a Heavy Phase Lock Gel tube (Eppendorf) containing 300 ul chloroform. The tube was inverted for 1 minute, and again centrifuged at max speed. The aqueous layer was then transferred to a mixture of 300 ul isopropanol and 300 ul high salt solution (0.8 M Na citrate, 1.2 M NaCl) for overnight precipitation. Following precipitation, RNA was purified using the standard protocol of the Qiagen RNeasy kit. To improve sequencing results, ribosomal RNA was removed from samples using the Ribo-Zero Bacterial rRNA Removal Kit (Illumina).

**RNA-seq library prep, sequencing, and normalization.** Construction of a cDNA library for Illumina sequencing was performed using the NEBNext Ultra RNA Library Prep Kit for Illumina. Final library concentration was quantified via KappaQuant (Kapa Biosystems). Paired-end RNA sequencing was completed at the Northwest Genomics Center at the University of Washington using a NextSeq 500 platform. After alignment to the H37Rv reference genome, normalization was completed for total read depth variation between samples, as well as gene length, using the R packages edgeR, limma, and lowess. Data are therefore presented in reads per kilobase of transcript per million mapped reads (RPKM).

**Rv1985c macrophage time course.** H37Rv:Rv1985c was grown to early log phase and split into standard growth media and media containing 100 ng/ml ATc one day before infection. The H37Rv:Δ1985c strain was cultured under standard conditions as well and kept in early log phase growth until the time of infection. Differentiated THP-1 cells were seeded into T25 flasks as described in chapter 1. For

infection, the early log phase H37Rv:Rv1985c -ATc, H37Rv:Δ1985c, and H37Rv:Rv1985c +ATc cultures were re-suspended in RPMI + 10% FBS at a concentration of 2.5 x $10^6$ bacteria/ml. 2 ml of each dilution (multiplicity of infection of 1) were then aliquoted to the appropriate number of flasks and allowed to infect at 37 $^oC$ for 4-6 hours. Following infection, each flask was aspirated of media, washed thoroughly with warm PBS three times to remove any extracellular bacteria, and fed back with 5 ml RPMI + 10% FBS (+/- 100 ng/ml ATc). Cultures were returned to 37 $^oC$, 5% $CO_2$ for 5 days of macrophage culture. Each time point of the experiment consisted of 3 separate induced and/or uninduced T25 flasks, and samples were collected at day 0 immediately following infection and again after five days of intra-macrophage growth for bacterial burden enumeration by plating.

## Results

**Hypoxia TRIP screen.** To pair with our macrophage regulon data, we additionally sought to investigate TRIP mutants within hypoxia, a stress associated with the formation of granulomas during chronic infection.

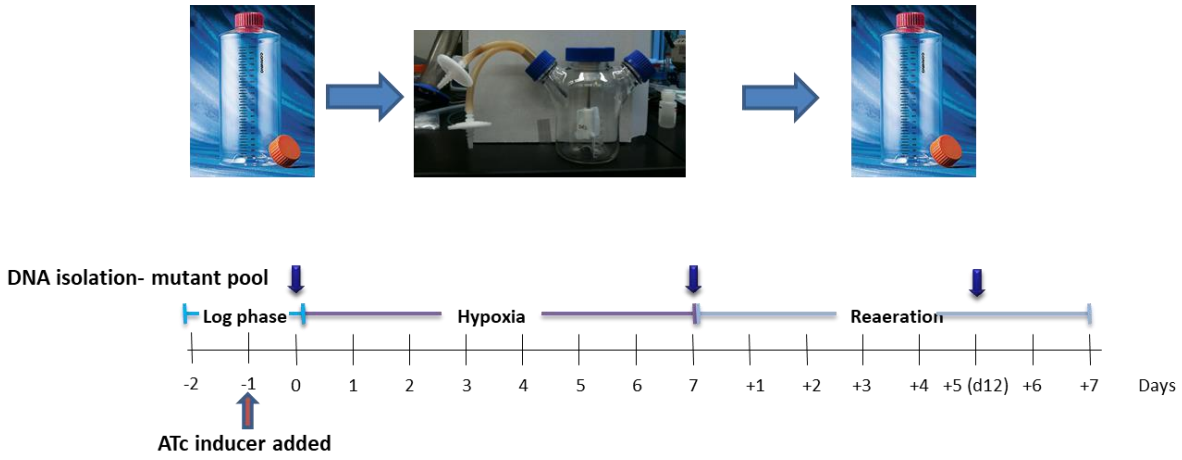



Figure 5. Overview of hypoxic TRIP screen. DNA isolation and sequencing were performed at day 0, 7, and 12 to assess the prevalence of mutants before, during, and after recovering from hypoxia.

The TRIP mutant pool was cultured under standard aerated conditions to early log phase growth, split, and one half induced with 100 ng/ml ATc one day before the induction of hypoxia. Before creating

hypoxic conditions, the cultures were again diluted to obtain an OD 0.1 at the outset of hypoxia. To induce hypoxia, cultures were transferred to airtight flasks, subjected to low oxygen gas flow with nitrogen balance, and kept under constant agitation to achieve 0.2% oxygen (Figure 5). The induced and uninduced cultures were kept under hypoxic conditions for seven days before being returned to aerated growth for an additional 5 days. For the initial screen, DNA sample collection was performed at three times: before creating hypoxic conditions, after 7 days hypoxic culture, and finally after 5 days of reaeration. The screen was performed on three separate biological replicates of the TRIP pool. Under hypoxic conditions, MTB enter a non-replicative state and little growth will occur in the bacterial population. For this reason, we measured for a phenotypic defect over the course of reaeration following hypoxic conditions. The day 12 time point, collected after 5 days aerated growth following hypoxia, was therefore compared to our log-phase data set outlined in the macrophage chapter. Both experiments measured relative abundance after 5 days of growth beginning from an OD 0.1 in uninduced and induced conditions. We hypothesized the induced TRIP strains recovering from exposure to hypoxia would accrue greater defects relative to the strains only induced under log phase conditions. 5 strains were identified with at least a two-fold greater defect as a result of hypoxia and reaeration relative to standard log-phase induction (Figure 6). 29 strains possessed similar defects in both conditions.

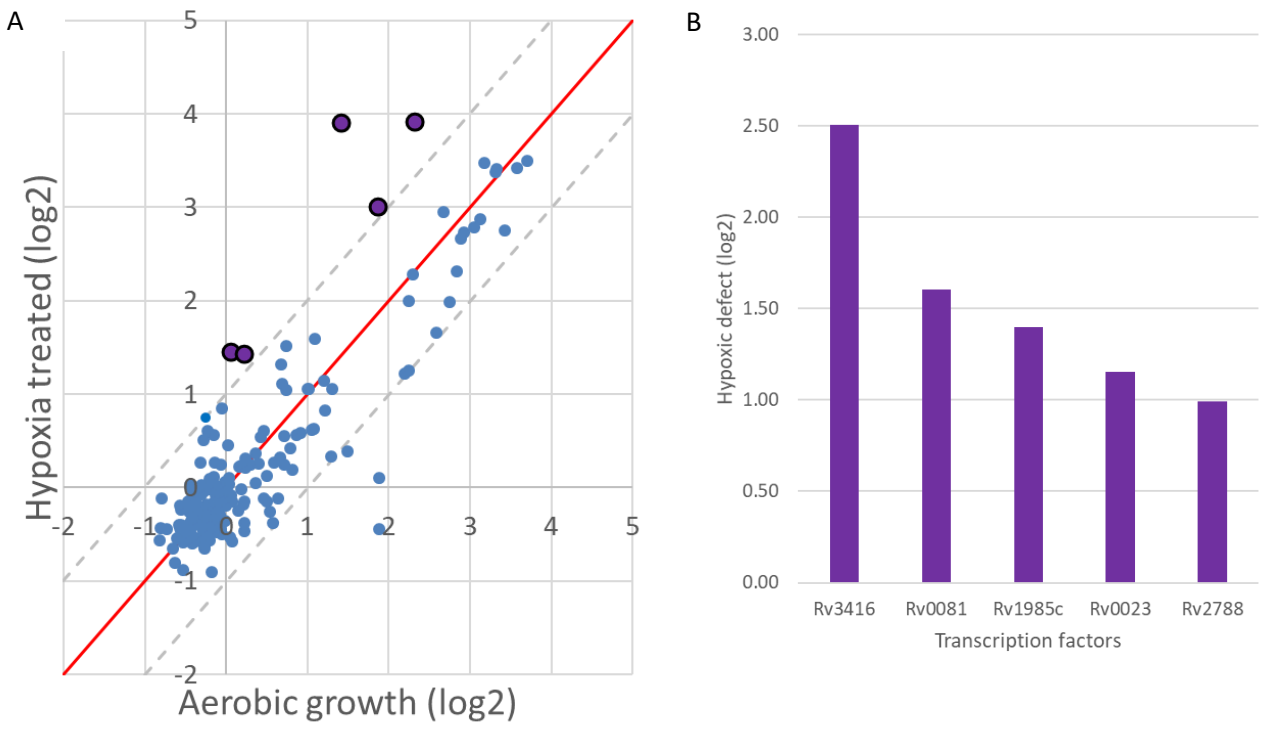



Figure 6. Comparison of TRIP hypoxia/reaeration phenotypes to log phase phenotypes. a. The day 12 hypoxia/reaeration uninduced-induced phenotypes were compared to the day 5 baseline log phase uninduced-induced phenotypes. The hypoxic screen was completed in triplicate and each marker represents an individual TRIP strain. b. Quantification of the 5 transcription factors with a greater than two-fold defect resulting from hypoxia and reaeration. All values are in log2 reads per million hits and the 5 phenotypic strains are highlighted in purple.

**Critical Oxygen Response Regulators.** The *Mycobacterium tuberculosis* hypoxic response has previously been well characterized by the Sherman Lab (Park 2003, Roberts 2004, Rustad 2008, Sherrid 2010). Over the course of hypoxia and reaeration, the expression of nearly a quarter of the MTB genome changes. While the immediate DosR response has been well characterized, little is known about the regulatory mechanisms driving the hundreds of genes participating in the enduring hypoxic response or reaeration response. To assess how well the regulons identified by the hypoxic TRIP screen aligned with the known MTB hypoxic response we compared the TRIP expression and ChIP data sets[46,47] for the 5

transcription factors with a hypoxia/reaeration defect to those genes induced during the enduring hypoxic response (Figure 7).

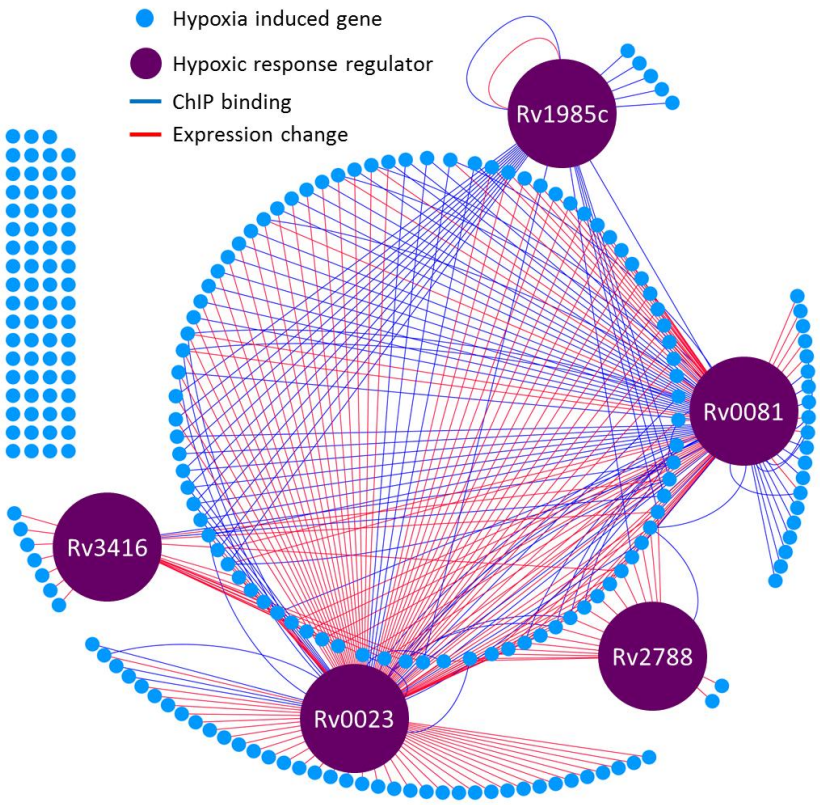



Figure 7. Comparison of TRIP log-phase expression regulons and ChIP data to the enduring hypoxic response. The five transcription factors (purple) are connected to the induced genes (blue) via expression changes (red lines) or ChIP binding peaks (blue). All interactions presented here fall below a p value cutoff of < 0.01. See references 46 and 47 for further detail.

239 genes are induced in the hypoxic response, 69% of which are either transcriptionally controlled by or have promoter binding sites for at least one of the five hypoxia and reaeration regulons. Considering this high degree of overlap, we have named the set of five transcription factors the Critical Oxygen Response Regulators (CORRs). Three of these transcription factors also were identified in our macrophage screen. Rv0023 and Rv0081 are extensively regulated in log-phase and bind many sites on the bacterial chromosome (Chapter 1, Table 1). Rv0023 induces 489 genes, represses 404, and has 241 ChIP binding peaks, while Rv0081 induces 283, represses 119, and possesses a ChIP binding profile of 647 peaks. Within the hypoxia induced gene network (Figure 7), Rv0023 regulates 102 genes by

expression and binds to 29 via ChIP. Rv0081 drives 53 expression changes and binds 63 genes by ChIP. In contrast, the other macrophage and hypoxia overlap regulon, Rv1985c, induces only 1 gene and represses 3 during log-phase growth. However, its ChIP binding profile includes 231 peaks. Unsurprisingly, in the CORR network all 23 putative Rv1985c regulatory connections to the enduring hypoxic response data set are via ChIP binding.

**Rv1985c.** Rv1985c is a putative LysR-type transcription factor with orthologs in *Escherichia coli* and *Corynebacterium glutamicum* that regulate the arginine operon and a lysine transporter, respectively[72,73]. However, relatively little work has been done outside of these cursory functional studies in the orthologs, and no studies have been undertaken on the role of this MTB regulator within either of the infection models presented in this thesis. A hallmark of LysR-type transcriptional regulators is their ability to bind target regions within the genome but exert no regulatory effect until an activating cofactor is present[74]. Given the high number of specific ChIP peaks that result in few transcriptional changes, we hypothesized that a specific condition within the macrophage and hypoxia/reaeration environments activates additional regulatory capacity of Rv1985c that has yet to be defined. This phenotype therefore presented to opportunity to define a novel regulon key to the adaptations to hypoxia/reaeration and potentially the macrophage. Accordingly, we next focused on better defining the Rv1985c phenotype and regulon within the hypoxic model.

**Rv1985c phenotype confirmation and bacterial burden.** To confirm our H37Rv:Rv1985c hypoxic phenotype and quantify changes in the bacterial burden over the course of hypoxia and reaeration, we repeated the hypoxic time course in triplicate and with additional time points (Figure 8).

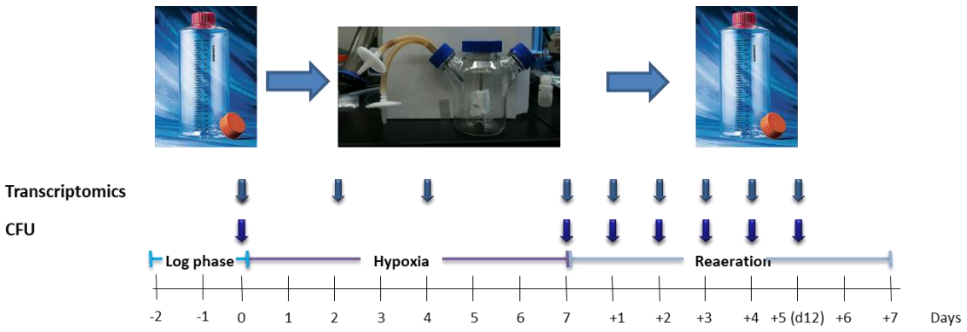



Figure 8. Expanded hypoxia/reaeration time course measurements to characterize changes in bacterial burden and the transcriptome. Rv1985c was individually repeated in triplicate flasks with a four flask as an uninduced control. In addition to the previous day 0, 7, and 12 time points, further hypoxic measurements were taken at days 2 and 4 and each day of reaeration was sampled.

Bacterial survival was determined by CFU at Days 0, 7 (end of hypoxia), and 8-12 (reaeration). Expression data was determined by RNA-seq at Day 0, 2, 4, 7 (hypoxia), and 8-12 (reaeration). In wild-type, Rv1985c is expressed at low levels during log phase growth, undergoes rapid induction to high levels during hypoxia, and immediately returns to baseline by the first day of reaeration. Given that we are inducing Rv1985c during the TRIP screen, we hypothesized the induction of Rv1985c during hypoxia would have little effect on the bacterial burden relative to wild-type, but the continued expression of Rv1985c above baseline during reaeration would result in decreased bacterial burden as reaeration progressed.

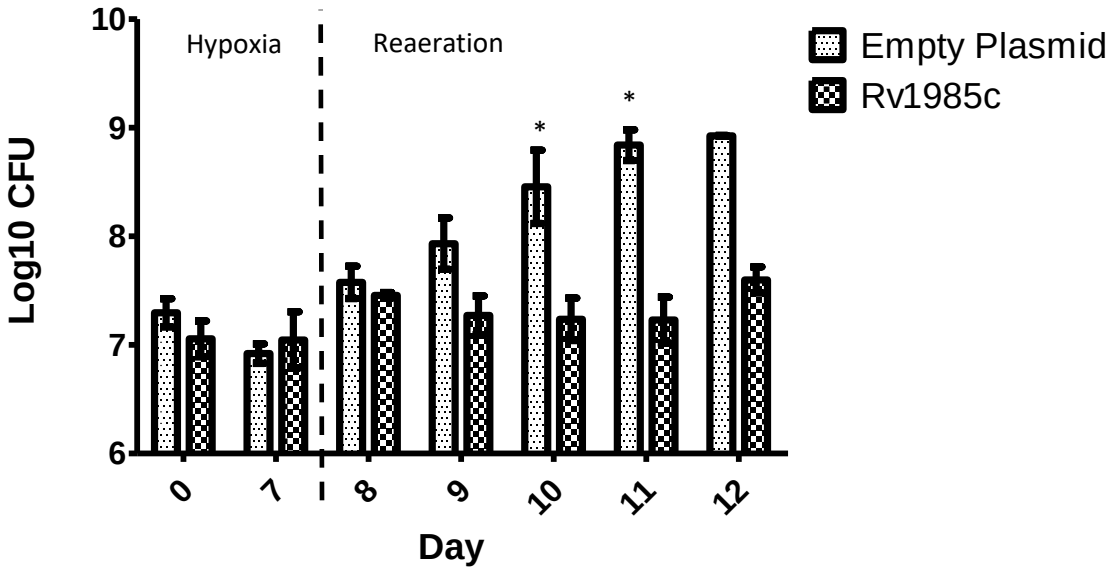



Figure 9. Bacterial burden of Rv1985c over the course of hypoxia and reaeration. Each strain was measured in triplicate biological replicates(excluding day 12) and values are graphed with standard error bars. Significance determined by student's t-test, * indicates $p<0.05$.

Both the empty plasmid control and Rv1985c induction mutant experienced negligible changes in bacterial survival over the course of hypoxia. The difference between the two strains remains non-significant at day 8, the first day of reaeration, and by day 9 divergent growth emerges as the control re-enters log phase growth and the mutant remains static. By day 10, three days of reaerated growth has produced a significant defect in the recovery of the Rv1985c induction mutant relative to the control strain (Figure 9, t-test $p<0.05$), totaling well over a log difference in bacterial burden. At day 5 of reaeration, the induction mutant exhibited its first uptick in growth, but the difference in bacterial burden remained roughly one and a half log difference. The defect due to Rv1985c induction previously observed in the TRIP screen therefore repeated, supporting the hypothesis of a defect in reaeration.

**H37Rv: Δ1985c mutant.** In addition to further characterizing the Rv1985c mutant individually, we also sought to identify any phenotype within the macrophage or hypoxia associated with a knockout of Rv1985c. Briefly, Rv1985c was knocked out via homologous recombination and replaced with a

hygromycin resistance gene and subjected to the same CFU and RNA isolation time course as outlined in Figure 8. Given that Rv1985c is strongly induced during hypoxia and quickly returns to baseline in reaeration, we hypothesized a defect would occur in the knockout during the seven-day hypoxic culture and recover in a manner similar to wild-type once reaeration occurred given that Rv1985c expression would again be negligible.

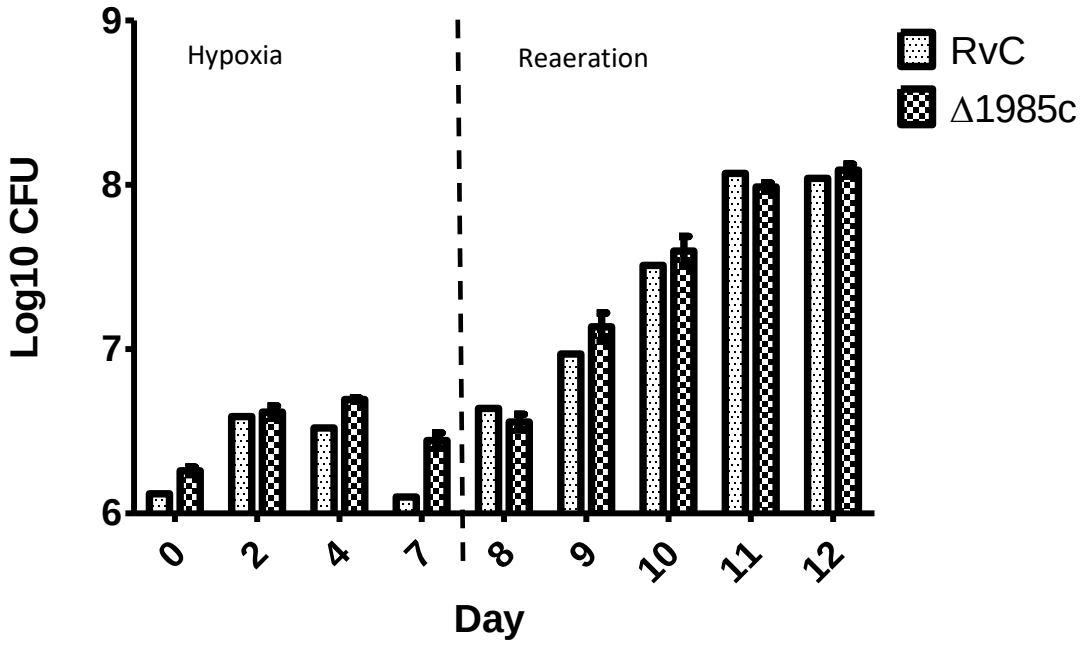



Figure 10. Bacterial burden of Δ1985c over the course of hypoxia and reaeration. The knockout strain was completed with triplicate biological replicates and standard error bars are shown.

The H37Rv wild-type strain and H37Rv:Δ1985c did not differ significantly over the course of the experiment (Figure 10). We expanded our bacterial burden characterization to include time points at days 2 and 4 of hypoxia, but there was little difference between the two strains. Over the full seven days of hypoxia, a minor amount of fluctuation in bacterial burden was observed totaling less than half a log, not an uncommon amount of variation in this type of experiment. Following reaeration, both strains lagged for one day as expected before re-entering log phase growth. Over the course of reaeration, no significant differences occurred in growth rate or total bacterial burden. Contrary to our

hypothesis, we did not find a hypoxic phenotype for H37Rv:Δ1985c, potentially indicating a redundant regulatory mechanism that can compensate for Rv1985c's absence. In addition to testing the Δ1985c mutant in hypoxia, we also determined if a phenotype occurred in our macrophage infection model. The standard five-day course of infection outlined in chapter 1 was performed, and bacterial burden determined by CFU.

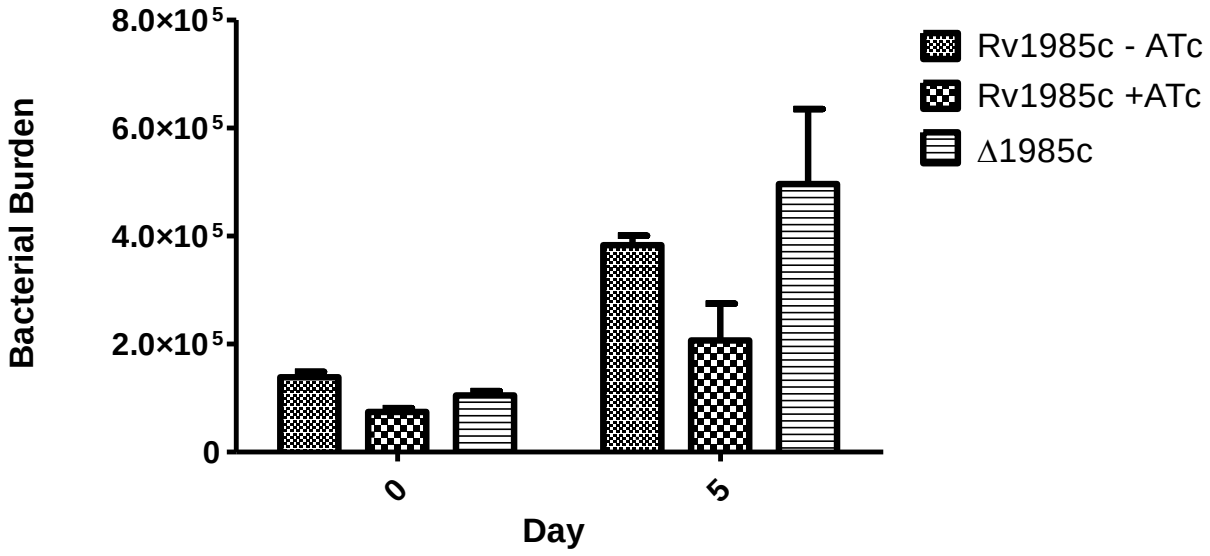



Figure 11. Macrophage infection phenotypes for Rv1985c and Δ1985c. The run was completed in triplicate biological measurements and standard error bars are shown. Differences trended towards significance by t-test but $p > 0.05$.

As expected, less growth was observed in the Rv1985c induced mutant in comparison to the uninduced control, with the induction mutant at roughly half the bacterial burden of the control (Figure 11). This confirmed the relative abundance phenotype observed in the chapter 1 macrophage screen. As with the hypoxia/reaeration characterization, the Δ1985c strain did not exhibit a growth defect in the macrophage infection model. In fact, the knockout exhibited more growth within the macrophage environment than did the wild-type control. Thus, in neither context in which the induction of Rv1985c produced a growth defect did the knockout of this same regulator impact growth.

**Transcriptomics to derive mechanistic hypothesis.** To this point, our screen data indicate the Rv1985c induction mutant carries a defect over the course of 5 days macrophage infection, and during the process of reaeration following 7 days of hypoxic culture. Individual strain follow-up confirmed these phenotypes, and provided the opportunity to collect RNA (Figure 8). We sought to define the transcriptional profile of each mutant over the course of hypoxia and reaeration to identify potential molecular mechanisms driving the phenotype. Recall, Rv1985c mediates only 4 expression changes during log-phase growth despite its widespread binding in the ChIP data set. Here, we hypothesized that novel, environment-specific transcriptional regulatory events would drive the reaeration phenotype.

Figure 12 provides a broad categorization of the fold change differential expression between the Rv1985c induction mutant and the empty plasmid wild-type control over the course of hypoxia and reaeration. Note that any genes not varying at least 2-fold between Rv1985c and the control at any point during the experiment were excluded. At day 0, there are minimal differences between the transcriptomes as expected. Over the course of the day 2, 4, and 7 hypoxia time points there is an increase in the number of differentially expressed genes, but both the number of genes and the magnitude of differential expression is relatively low. This is not surprising given that no phenotype was observed following Rv1985c induction during hypoxia. At day 8, the first day of reaeration, the number of differentially expressed genes is again quite low – it appears as if the two transcriptomes converged back towards log phase similarity. Euclidian clustering analysis showed days 0 and 8 were the two most similar in terms of fold change differences across the transcriptome. As reaeration progresses, the number of divergently expressed genes quickly increases. Furthermore, the magnitude of these expression differences is increased. Gene ontology analysis of the divergently regulated gene sets from days 9-12 failed to identify enriched associations within the biological processes or molecular functions

categories – potentially indicating that they hypoxia defect of the Rv1985c induction mutant is not due to dysregulation of one particular stress response but is multifactorial.

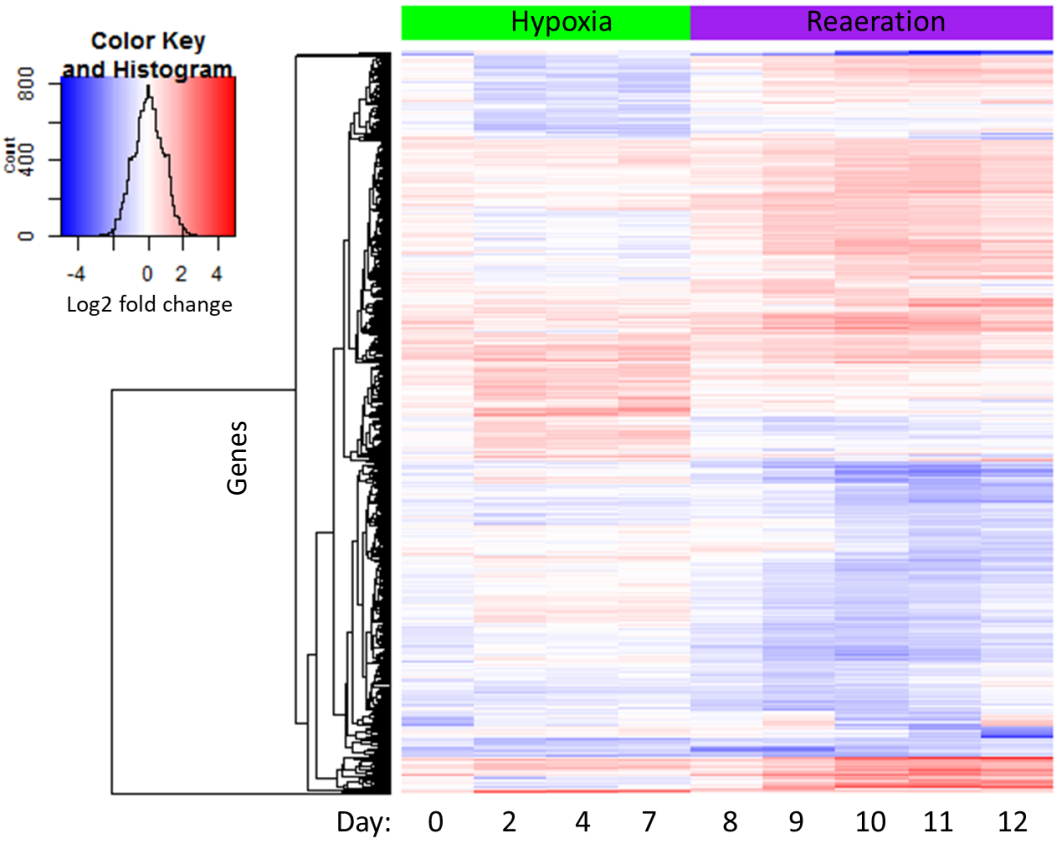



Figure 12. Rv1985c induction vs empty plasmid differential expression analysis. Samples were collected in triplicate and data are graphed as fold change log2 reads per kilobase of transcript per million mapped reads (RPKM). Those genes that did not vary at least 2-fold over the course of the experiment are excluded. Relative to Rv1985c, blue indicates repression and red induction. Color breaks are drawn to highlight expression changes beginning at two-fold and reaching maximum intensity at four-fold or above.

Recall the Rv1985c induction strain begins to accrue its bacterial growth defect by the second day of reaeration (Figure 9), when we see the widespread divergent expression differences build. The growth defect by CFU reaches significance on Day 10, when we correspondingly see nearly a quarter of the genome has become differentially expressed between Rv1985c and the control (Figure 12). Based upon this, we hypothesized a group of genes are mis-regulated in the Rv1985c induction mutant at day 8, and this drives the failure to reaerate properly and widespread dysregulation throughout the rest of the

transcriptome. We therefore examined the largest expression differences at Day 8 more closely, just before the widespread separation of the two transcriptomes occurs.

**The ribosomal protein locus.** To inform a hypothesis regarding the molecular mechanism driving the Rv1985c induction phenotype, we next examined the specific genes with a 4-fold or greater expression difference at Day 8 (Table 2).

Table 2. Genes with at least 4-fold differential expression between Rv1985c and the empty plasmid on day 1 of reaeration. Negative log2 values in the Empty-Ind column indicate genes that are induced in Rv1985c, while positive values indicate repression.

| Gene | NAME | PRODUCT | Empty-Ind (log2) |
|---|---|---|---|
| Rv1985c | | transcriptional regulator, lysR-family | -7.9 |
| Rv1316c | ogt | methylated-DNA-protein-cysteine methyltransferase ogt | -2.5 |
| Rv1317c | alkA | ada regulatory protein alkA | -2.1 |
| Rv1318c | | adenylate cyclase | -2.1 |
| Rv0721 | rpsE | 30S ribosomal protein S5 rpsE | 2.2 |
| Rv0723 | rplO | 50S ribosomal protein L15 rplO | 2.2 |
| Rv0719 | rplF | 50S ribosomal protein L6 rplF | 2.3 |
| Rv3477 | PE31 | PE family protein | 2.3 |
| Rv0720 | rplR | 50S ribosomal protein L18 rplR | 2.3 |
| Rv0722 | rpmD | 50S ribosomal protein L30 rpmD | 2.3 |
| Rv0714 | rplN | 50S ribosomal protein L14 rplN | 2.6 |
| Rv0717 | rpsN1 | 30S ribosomal protein S14 rpsN1 | 2.6 |
| Rv0718 | rpsH | 30S ribosomal protein S8 rpsH | 2.6 |
| Rv0715 | rplX | 50S ribosomal protein L24 rplX | 2.7 |
| Rv0716 | rplE | 50S ribosomal protein L5 rplE | 2.7 |

Unsurprisingly, Rv1985c is highly induced in relation to all other genes. The three remaining induced genes appear to be part of an operon. ogt and alkA both function within DNA repair, and adenylate

cyclase catalyzes the formation of the important signaling molecule cyclic AMP. Strikingly, of the 11 repressed genes in Rv1985c, 10 of them are ribosomal proteins. Furthermore, these ribosomal proteins compose a continuous operon, from Rv0714 – Rv0723. *Mycobacterium tuberculosis* possesses ~50 ribosomal proteins, and in our log-phase TRIP expression data very few ribosomal protein genes were transcriptionally altered by the induction of any TRIP mutant. Given the high number of ribosomal genes repressed by Rv1985c during reaeration, we sought to more thoroughly examine this locus over the course of hypoxia and reaeration.

In wild-type, the ribosomal protein genes are transcriptionally repressed during hypoxia (days 2, 4, 7) and immediately increase expression upon reaeration. Figure 13 compares the fold change expression difference between H37Rv:Rv1985c and the empty plasmid control in this ribosomal protein locus.

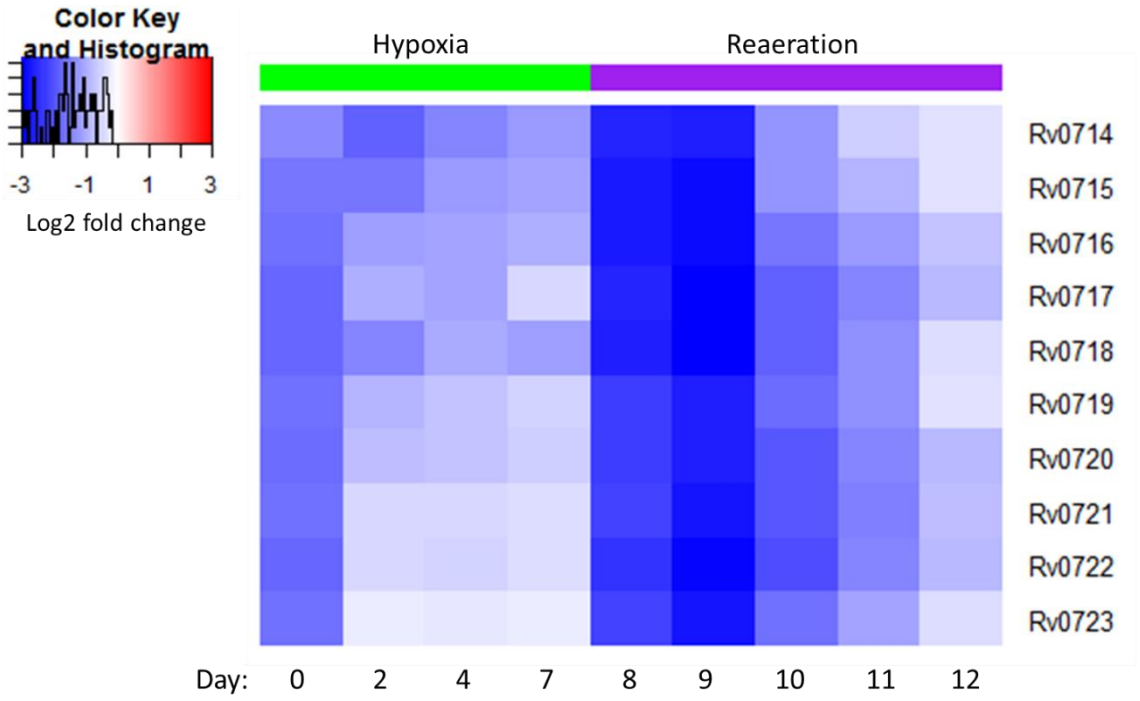



Figure 13. Ribosomal protein fold change expression heat map for Rv1985c induction vs wild-type control over the course of hypoxia and reaeration. Relative to Rv1985c, blue shading indicates repression, red induction. The heat map represents the magnitude of fold change log2 RPKM data from three biological replicates. Color breaks begin at 2-fold and progress to solid colors at 8-fold.

Fold change comparisons reveal that at day 0, there is a roughly 2-fold repression of this ribosomal protein locus due to the addition of ATc one day before the experiment. Throughout hypoxia, there are minimal fold change differences between H37Rv:Rv1985c and the control, though the first 5 genes in the operon may be slightly more repressed by the induction mutant. At day 8, where we saw minimal expression differences elsewhere in the transcriptome (Figure 12), the ribosomal proteins are uniformly repressed, some to nearly an 8-fold difference. This strong repression continues through days 9 and 10 (the 2nd and 3rd of reaeration), where we see the bacterial burden differences take effect (Figure 9). At four days of reaeration, the repression begins to lessen, and by the fifth day of reaeration the locus closely resembles wild-type with only minor repression in some cases.

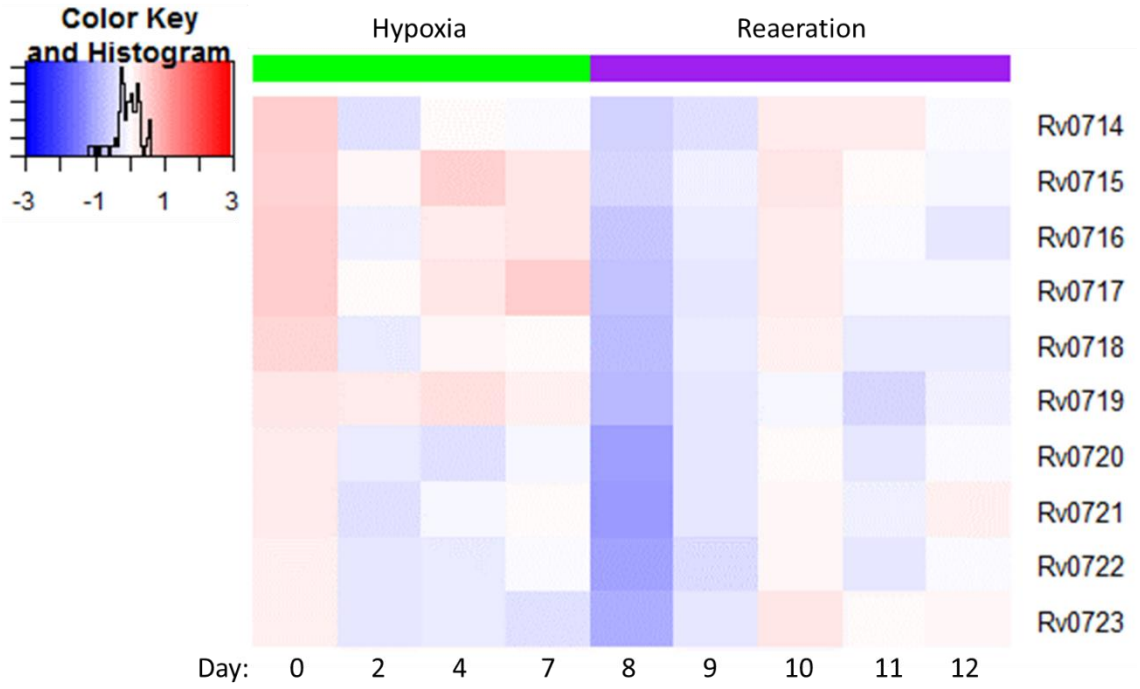



Figure 14. Ribosomal protein fold change expression heat map for Δ1985c induction vs wild-type control over the course of hypoxia and reaeration. Relative to Δ1985c, blue shading indicates repression, red induction. The heat map represents the magnitude of fold change log2 RPKM data from three biological replicates. The color breaks here are identical to those in Figure 13.

If Rv1985c was responsible for repressing the ribosomal protein locus during hypoxia, we hypothesized the deletion of Rv1985c would prevent this from occurring. However, the Δ1985c expression data closely mirrors the wild-type across the ribosomal protein locus for both hypoxia and reaeration, as evidenced by minimal fold-change differences rarely exceeding 2-fold in the figure above (Figure 14). Considering our CFU data demonstrated no phenotype during hypoxic culture for the knockout, this was not entirely surprising and potentially indicates redundant regulatory mechanisms capable of acting upon the ribosomal protein locus.

**Discussion**

Here we set out to define the regulators and regulons of MTB critical to the orchestration of the hypoxia/reaeration response. Past efforts had defined the landscape of transcriptional changes throughout hypoxia, but little was understood regarding how specific regulators orchestrated the response through specific, evolved subnetworks. Furthermore, it remains an open question within the field as to how MTB's adaptations influence or derive from known correlates of disease progression in the granuloma.

Here, we identified 5 transcription factors with an at least 2-fold growth defect relative to wild-type control over the course of hypoxia and reaeration (Figure 6). Subsequently, we sought to understand how well these five regulators mapped to the known genes of the hypoxic response. 70% of the enduring hypoxic response induced genes mapped to the regulons of the five TFs via known expression changes or DNA-binding locations (Figure 7). In light of this overlap, we have named the set of TFs the Critical Oxygen Response Regulators. Rv0081 is a member of the DosR regulon, and potentially forms a hierarchical link between the regulation of the immediate DosR response and the enduring hypoxic response. In log phase, Rv0081 modulates a large portion of the transcriptome (283 induced genes, 119

repressed) and displays widespread DNA binding (647 ChIP peaks). Correspondingly, this TF overlaps with the enduring hypoxic response via 53 expression changes and 63 ChIP binding events. Such widespread expression changes and binding events may indicate a broader stress response pathway that is regulated by Rv0081 as a part of the hypoxic response. Similarly, Rv0023 exerts widespread change in the log-phase transcriptome (489 induced, 404 repressed) and binds 241 locations via ChIP. 102 of its expression changes and 29 DNA-binding locations overlap with the hypoxia induced genes. There is no known functional characterization for this TF. In log-phase Rv3416 mediates 255 induction events, and 116 repression events, along with 21 ChIP binding peaks. It possesses 24 connections to the enduring hypoxic response, 2 by ChIP and 22 by expression. Rv3416 is a known antigen expressed by MTB during macrophage infection and is associated with the downregulation of the apoptotic pathway within macrophages[75-77]. Despite this finding, this TF was not identified by our macrophage screen, perhaps due to the lack of an *in-vivo* context for infection progression. Rv2788 possesses a relatively sparse regulon compared to the other CORRs above, inducing 9 genes and repressing 72 in log-phase, along with 5 ChIP binding peaks. On the CORR map, it connects to the hypoxia induced genes via 18 expression changes and 1 DNA-binding peak. Rv2788 is in fact a manganese dependent transcriptional repressor of manganese transporters, and has been shown to be essential for adaptation to manganese-rich conditions[78]. Further studies in *M. smegmatis* demonstrated its overexpression improved survival to a variety of stresses including hydrogen peroxide, acid stress, and antibiotic treatment[79].

It should be noted that 3 of these CORRs (Rv0081, Rv0023, and Rv1985c) were also identified by our macrophage phenotype screen. Potentially this indicates part, or all, of the hypoxic response network is critical to the adaption within macrophages as well as hypoxia, or that these CORRs adapt subsets of general stress response mechanisms to a specific context.

We acknowledge the high number of binding and transcriptional changes attributed to these five transcription factors, and that not all of them overlap with the enduring hypoxic response (Figure 7). It is possible additional stress contexts exist for which these transcription factors respond, or that specific cofactors were not present to activate certain genes. Furthermore, DNA-binding proteins have been shown to carry out functions in addition to transcription factor activity, including that of a nucleoid-associated protein responsible for higher-order chromosomal structure. In MTB, EspR has been published as both a transcription factor and a nucleoid-associated protein[81,82].

Review of our CORR regulons identified Rv1985c as an intriguing candidate for further characterization. Rv1985c is a LysR-type transcription factor known to mediate only 4 expression changes in log-phase, despite a relatively widespread ChIP binding profile of 231 locations in the genome. Its connection to the hypoxia-induced genes was mediated by 23 DNA-binding events. A hallmark of LysR-type transcription factors is their ability to bind to DNA target sequences without exerting a regulatory effect until the secondary binding of a cofactor. This uniquely positions these TFs to rapidly adapt an organism to a given stress, and fits with our observed widespread, quiescent DNA binding. Accordingly, we hypothesized this TF could possess a hypoxia/reaeration specific regulon, and these novel expression changes would then drive the observed growth defects in our screens. Accordingly, we carried out individual follow-up of Rv1985c via bacterial burden studies and RNA-seq.

We confirmed the presence of a growth defect for Rv1985c (Figure 9) and clarified that it did not occur during hypoxia itself (when Rv1985c is induced in the wild-type), but rather the Rv1985c induction mutant lags during reaeration and does not appear able to resume log-phase growth. This led to the hypothesis that the continued expression of Rv1985c during reaeration, when wild-type expression levels return to baseline, would drive the continued expression of a regulon responsible for the

phenotype. RNA-seq identified relatively few differences between the Rv1985c induction strain and empty plasmid control in log-phase and hypoxia (Figure 12). The first day of reaeration also produced few expression differences, and was quite similar to the log-phase comparison between the two. Beginning at day 9, widespread transcriptional divergence between the mutant and control were observed, with the magnitude of divergent regulation increasing as reaeration progressed. On days 3 and 4 of reaeration, nearly 1,000 genes (~1/4 of the MTB genome) were disparately regulated. This rapid expression divergence corresponded with the onset of a growth phenotype between the two strains. Given that Rv1985c is continually induced throughout the experiment, we hypothesized the responsible differential regulon must have continued to be expressed after hypoxia into the first day of reaeration and could be identified there before the widespread transcriptional divergence occurred. Table 2 identified a continuous operon of ribosomal proteins are strongly downregulated by Rv1985c at day 8. Analysis of their expression over the course of hypoxia and reaeration revealed these ribosomal proteins are downregulated to a similar level as in H37Rv:Rv1985c during hypoxia, but Rv1985c induction is sufficient to drive their continued repression throughout reaeration. Interestingly, the Δ1985c strain did not ablate ribosomal protein repression, potentially indicating redundant regulatory mechanisms for these proteins' transcription. Therefore, our current working hypothesis is that the induction of Rv1985c is sufficient to continuously repress the transcription of the identified ribosomal protein locus, and this in turn reduces ribosome availability in the reaeration response, preventing the bacteria from resuming normal growth. This identification represents a novel regulatory mechanism for ribosomal proteins in our MTB data sets, as we did not identify regulators associated with control of any of the 50 total ribosomal proteins in our log-phase networks.

In the experiments depicted by figures 10 and 11, we tested whether a chromosomal knockout of Rv1985c would possess a defect in either hypoxia/reaeration or the macrophage. Given that Rv1985c is

strongly induced in response to hypoxia in the wild-type, we did hypothesize the knockout would carry a hypoxic defect. In both macrophages and the hypoxia/reaeration time course the Δ1985c phenotype was not significantly different from the control. This illustrates the potential importance of the regulons identified in both our macrophage and hypoxia screens. While it is possible to illuminate gene regulatory networks via gene disruption of the regulators, we hypothesized there were additional regulators previous screens had missed, and that induction mutants would better preserve the networks and elucidate more subtle, novel phenotypes. The lack of a phenotype in the knockout in both screens in which Rv1985c induction drove a phenotype confirms this point and emphasizes the contributions these regulons could potentially make to the field moving forward.

**Future Work.** An important biological question we have yet to answer in regard to Rv1985c is whether the downregulation of the identified ribosomal proteins directly leads to reduced numbers or different types of ribosomes. Direct quantification of ribosomes during reaeration in the Rv1985c induction mutant is required to confirm the validity of our current hypothesis. More specifically, the ribosomal proteins identified in Table 2 correspond both to different subunits and sites of protein synthesis within the ribosomes. If the mechanism is not simply a lower number of total ribosomes, examination of the availability of specific ribosomal subunits and protein translation efficiency and output may yield the specific cause of the Rv1985c induction defect.

Additionally, we hypothesized the widespread ChIP-binding of Rv1985c was evidence of a context-specific regulon. This was not necessarily the case with the ribosomal proteins, as RNA-seq revealed they are slightly repressed during log-phase growth in the Rv1985c induction mutant. Therefore, repression of these ribosomal proteins could be achieved without the hypoxic context – though it was to a lesser degree. The *E. coli* ortholog of Rv1985c is predicted to bind arginine and lysine as cofactors, and

a recent paper identified limited expression changes mediated by Rv1985c were dependent on lysine[80,73]. Experiments investigating whether the depletion or excess of the potential cofactors mediated gene expression changes should be undertaken. Potentially, this could explain the discrepancy between the magnitude of ribosomal protein repression between Rv1985c induction during log phase and hypoxia. Additionally, there is the potential for a cofactor specific regulon to exist separate from the ribosomal protein mechanism hypothesized here.

The failure of the knockout Rv1985c strain to ablate repression of the ribosomal protein locus implies the existence of a hierarchy of regulation that can exert redundant function should one regulator be lost. We proposed to clone the ribosomal protein locus of interest into our induction plasmid, and transform this into the H37Rv:Δ1985c strain. Cross-linking followed by targeted pulldown of the ribosomal protein locus in the absence of Rv1985c and sequencing of any bound regulators may elucidate other regulators responsible for this key locus. It is possible the potential cofactors discussed above could also contribute to the putative hierarchy of regulation. These pulldowns should also be completed in the presence or absence of arginine and lysine in order to determine if cofactors contribute to which regulator preferentially binds and exerts function.

Finally, further work should be undertaken to characterize the degree to which the CORR regulons functionally organize the hypoxic response. Our current network of overlap with the enduring hypoxic response is predicated on log-phase data. Transcriptomic analysis of each of the CORRs should be undertaken during hypoxia and reaeration and mapped again to the hypoxia induced genes. This could also potentially identify novel regulons not included in the enduring hypoxic response that are critical to hypoxic adaptation.

**References**


1. Bañuls, AL, et al. *Mycobacterium tuberculosis: ecology and evolution of a human bacterium.* Journal of Medical Microbiology, 2015. 64: 1261-1269.

2. World Health Organization. Global Tuberculosis Report 2017.

3. Frieden, TR, et al. *Tuberculosis.* Lancet, 2003. 362: 887-899.

4. Russell, DG, Barry, CE, and Flynn, JL. *Tuberculosis: What we don't know can, and does, hurt us.* Science, 2010. 328: 852-856.

5. Gagneux, S and Small, PM. *Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development*. Lancet Infect. Dis. 2007. 7: 328-337.

6. Comas, I, et al. *Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans.* Nat Gen, 2013. 45: 1176-1182.

7. Roberts, CA, et al. *Letter to the editor: was tuberculosis present in Homo erectus in Turkey?* Am J Phys Anthropol, 2009. 139: 442-444.

8. Reibman, J. *Phthisis and the Arts, in Tuberculosis.* WN Rom and SM Garay, Editors. 1996. Little, Brown, and Co: Boston. 21-34.

9. Marrakchi, H, Laneelle, MA, and Daffe, M. *Mycolic Acids: Structures, Biosynthesis, and Beyond.* Cell Chemistry and Biology, 2014. 21: 67-85.

10. Pai, M, et al. *Tuberculosis.* Nature Reviews Disease Primers, 2016. 2: 1-23.

11. Dannenberg, AM Jr. *Immune mechanisms in the pathogenesis of pulmonary tuberculosis.* Review of Infectious Diseases, 1989. 11 (supp 2): S369-S378.

12. Via, LE, et al. *Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates.* Infect Immun, 2008. 76: 2333-2340.

13. Vynnycky, E and Fine, PE. *The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection.* Epidemiol Infect, 1997. 119: 183-201.

14. Barry, CE 3rd, et al. *The spectrum of latent tuberculosis: rethinking the biology and intervention strategies.* Nat Rev Microbiol, 2009. 7: 845-855.

15. Antonucci, G, et al. *Risk factors for tuberculosis in HIV-infected persons: a prospective cohort study.* JAMA, 1995. 274: 143-148.

16. Bean, AG, et al. *Structural deficiencies in granuloma formation in TNF gene-targeted mice underlie the heightened susceptibility to aerosol Mycobacterium tuberculosis infection, which is not compensated for by lymphotoxin.* J Immunol, 1999. 162: 3504-3511.

17. Chakravarty, SD, et al. *Tumor necrosis factor blockade in chronic murine tuberculosis enhances granulomatous inflammation and disorganizes granulomas in the lungs.* Infect Immun, 2008. 76: 916-926.

18. Vandal, OH, Nathand, CF, and Ehrt, S. *Acid resistance in mycobacterium tuberculosis.* J Bacteriol, 2009. 191: 4714-4721.

19. Huaman, MA, Deepe, GS Jr, and Fichtenbaum, CJ. *Elevated circulating concentrations of interferon-gamma in latent tuberculosis infection.* Pathog Immun, 2016. 1: 291-303.

20. Zewdie, M, et al. *Ex-vivo characterization of regulatory T cells in pulmonary tuberculosis patients, latently infected persons, and healthy endemic controls.* Tuberculosis, 2016. 100: 61-68.

21. Heidarnezhad, F, et al. *Evaluation of Interleukin 17 and Interleukin 23 expression in patients with active and latent tuberculosis infection.* Iran J Basic Med Sci, 2016. 19: 844-850.

22. Fox, GJ, et al. *Surgery as an adjunctive treatment for multidrug-resistant tuberculosis: an individual patient data metaanalysis.* Clin Infect Dis, 2016. 62: 887-895.

23. Daniels, M and Hill, AB. *Chemotherapy of pulmonary tuberculosis in young adults; an analysis of the combined results of three Medical Research Council trials.* Br Med J, 1952. 1: 1162-1168.

24. Dheda, K, et al. *Global control of tuberculosis: from extensively drug-resistant to untreatable tuberculosis.* Lancet Respir Med, 2014. 2: 321-338.

25. Cox, E and Laessig, K. *FDA approval of bedaquiline – the benefit-risk balance for drug-resistant tuberculosis.* N Eng J Med, 2014. 371: 689-691.

26. Mangtani, P, et al. *Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials.* Clin Infect Dis, 2014. 58: 470-480.

27. Roy, A, et al. *Effect of BCG vaccination against Mycobacterium tuberculosis infection in children: systematic review and meta-analysis.* BMJ, 2014. 349: g4643.

28. Trunz, BB, Fine, P, and Dye, C. *Effect of BCG vaccination on childhood tuberculosis meningitis and military tuberculosis worldwide: a meta-analysis and assessment of cost effectiveness.* Lancet, 2006. 367: 1173-1180.

29. Izzo, AA. *Tuberculosis vaccines – perspectives from the NIH/NIAID Mycobacteria vaccine testing program.* Curr Op Immun, 2017. 47: 78-84.

30. Tan, S, Russell, DG. *Trans-species communication in the Mycobacterium tuberculosis-infected macrophage*. Immunol Rev, 2015. 264: 233-248.

31. Schaible, UE, et al. *Cytokine activation leads to acidification and increases maturation of Mycobacterium avium-containing phagosomes in murine macrophages.* J Immunol, 1998. 160: 1290-1296.

32. Via, LE, et al. *Effects of cytokines on mycobacterial phagosome maturation.* J Cell Sci, 1998. 111: 897-905.

33. Stutz, MD, et al. *Mycobacterium tuberculosis: Rewiring host cell signaling to promote infection.* J Leukoc Biol, 2017. DOI: 10.1002/JLB.4MR0717-277R

34. Dorhoi, A, et al. *Type I IFN signaling triggers immunopathology in tuberculosis-susceptible mice by modulating lung phagocyte dynamics.* Eur J Immunol, 2014. 44: 2380-2393.

35. Manca, C, et al. *Hypervirulent M. tuberculosis W/Beijing strains upregulate type I IFNs and increase expression of negative regulators of the Jak-Stat pathway.* J Interferon Cytokine Res, 2005. 25: 694-701.

36. McNab, FW, et al. *Type I IFN induces IL-10 production in an IL-27-independent manner and blocks responsiveness to IFN-y for production of IL-12 and bacterial killing in Mycobacterium tuberculosis infected macrophages.* J Immunol, 2014. 193: 3600-3612.

37. Watson, RO, Manzanillo, PS, and Cox, JS. *Extracellular M. tuberculosis DNA targets bacteria for autophagy by activating the host DNA-sensing pathway.* Cell, 2012. 150: 803-815.

38. Castillo, EF, et al. *Autophagy protects against active tuberculosis by suppressing bacterial burden and inflammation.* Proc Natl Acad Sci, 2012. 109: E3168-E3176.

39. Divangahi, M, Behar, SM, and Remold, H. *Dying to live: how the death modality of the infected macrophage affects immunity to tuberculosis.* Adv Exp Med Biol, 2013. 783: 103-120.

40. Srinivasan, L, Ahlbrand, S, and Briken, V. *Interaction of Mycobacterium tuberculosis with host cell death pathways.* Cold Spring Harb Perspect Med, 2014. 4: a022459.

41. Behar, SM, et al. *Apoptosis is an innate defense function of macrophages against Mycobacterium tuberculosis.* Mucosal Immunol, 2011. 4: 279-287.

42. Schnappinger, D, et al. *Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment.* J Exper Med, 2003. 198: 693-704.

43. Rohde, KH, Abramovitch, RB, and Russell, DG. *Mycobacterium tuberculosis invasion of macrophages: linking bacterial gene expression and environmental cues.* Cell Host Microbe, 2007. 2: 352-364.

44. Rohde, KH, et al. *Linking the transcriptional profiles and the physiological states of Mycobacterium tuberculosis during an extended intracellular infection.* PLoS Pathog, 2012. 8: e1002769.

45. Homolka, S, et al. *Functional genetic diversity among Mycobacterium tuberculosis complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival.* PLoS Pathog, 2010. 6: e1000988.

46. Minch, KJ, et al. *The DNA-binding network of Mycobacterium tuberculosis.* Nat Commun, 2014. 6: 5829.

47. Rustad, TR, et al. *Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network.* Genome Biology, 2014. 15: 502.

48. Brugarolas, P, et al. *The oxidation-sensing regulator (MosR) is a new redox-dependent transcription factor in Mycobacterium tuberculosis.* Journ Biol Chem, 2012. 287: 37703-37712.

49. Kumar, N, et al. *Crystal structure of the transcriptional regulator Rv1219c of Mycobacterium tuberculosis.* Protein Sci, 2014. 23: 423-432.

50. Neyrolles, O, et al. *Mycobacteria, metals, and the macrophage.* Immunological Reviews, 2015. 264: 249-263.

51. Bhattacharya, M, and Kumar Das, A. *Inverted repeats in the promoter as an autoregulatory sequence for TcrX in Mycobacterium tuberculosis.* Biochem Biophys Res Comm, 2011. 415: 17-23.

52. Marcus, SA, et al. *CsoR is essential for maintaining copper homeostasis in Mycobacterium tuberculosis.* PLoS One, 2016. DOI:10.1371.

53. Wolschendorf, F, et al. *Copper resistance is essential for virulence of Mycobacterium tuberculosis.* Proc Natl Acad Sci USA, 2011. 108: 1621-1626.

54. Santangelo, M, et al. *Mce3R, a tetR-type transcriptional repressor, controls the expression of a regulon involved in lipid metabolism in Mycobacterium tuberculosis.* Microbiology, 2009. 155: 2245-2255.

55. Kim, MJ, et al. *Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism.* EMBO Mol Med, 2010. 2: 258-274.

56. Peyron, P, et al. *Foamy macrophages from tuberculosis patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence.* PLoS Path, 2008. DOI: 4:e1000204.

57. Patel, N, et al. *A novel 6-benzyl ether benzoxaborole is active against Mycobacterium tuberculosis in vitro.* Antimicrob Agents Chemother, 2017. 61: e01205-17.

58. Balganesh, M, et al. *Rv1218c, an ABC transporter of Mycobacterium tuberculosis with implications in drug discovery.* Antimicrob Agents Chemother, 2010. 54: 5167-5172.

59. Corbett, EL, et al. *The growing burden of tuberculosis: global trends and interactions with the HIV epidemic.* Arch Intern Med, 2003. 163: 1009-1021.

60. McShane, H. *Co-infection with HIV and TB: double trouble.* Int J STD AIDS, 2005. 16: 95-100.

61. Adler, JJ and Rose, DN. *Transmission and pathogenesis of tuberculosis*. In Tuberculosis (1996). Rom, WN and Garay, SM (ed). Boston: Little, Brown, pp. 129-140.

62. Kaplan, G, et al. *Mycobacterium tuberculosis growth at the cavity surface: a microenvironment with failed immunity.* Infect Immun, 2003. 71: 7099-7108.

63. Olender, S, et al. *Low prevalence and increased household clustering of Mycobacterium tuberculosis infection in high altitude villages in Peru.* Am J Trop Med Hyg, 2003. 68: 721-727.

64. Aly, S, et al. *Oxygen status of lung granulomas in Mycobacterium tuberculosis infected mice.* J Pathol, 2006. 210: 298-305.

65. Via, LE, et al. *Tuberculosis granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates.* Infect Immun, 2008. 76: 2333-2340.

66. Flynn, J. *Lessons from experimental Mycobacterium tuberculosis infections.* Microbes Infect, 2006. 8: 1179-1188.

67. Lin, PL, et al. *Early events in Mycobacterium tuberculosis infection of cynomolgus macaques.* Infect Immun, 2006. 74: 3790-3803.

68. Park, HD, et al. *Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis.* Mol Microb, 2003. 48: 833-843.

69. Roberts, DM, et al. *Two sensor kinases contribute to the hypoxic response of Mycobacterium tuberculosis.* J Biol Chem, 2004. 279: 23082-23087.

70. Kumar, A, et al. *Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor.* Proc Natl Acad Sci USA, 2007. 104: 11568-11573.

71. Rustad, TR, et al. *The enduring response of Mycobacterium tuberculosis.* PLoS One, 2008. 3: e1502.

72. Nandineni, MR and Gowrishankar, J. *Evidence for an arginine exporter encoded by yggA (argO) that is regulated by the LysR-type transcriptional regulator ArgP in Escherichia coli.* J Bacteriol, 2004. 186: 3539-3546.

73. Schneefeld, M, et al. *The transcriptional regulator LysG (Rv1985c) of Mycobacterium tuberculosis activates lysE (Rv1986) in a lysine-dependent manner.* PLoS One, 2017. 12: e0186505.

74. Maddocks, SE and Oyston, PC. *Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins.* Microbiology, 2008. 154 (Pt 12): 3609-3623.

75. Mehto, S, et al. *Mycobacterium tuberculosis and human immunodeficiency virus type 1 cooperatively modulate macrophage apoptosis via toll like receptor 2 and calcium homeostasis.* PLoS One, 2015. 10: e0131767.

76. Chadha, A, et al. *Suppressive role of neddylation in dendritic cells during Mycobacterium tuberculosis infection.* Tuberculosis (Edinb), 2015. 95: 599-607.

77. Selvakumar, A, et al. *Reciprocal regulation of reactive oxygen species and phosphor-CREB regulates voltage gated calcium channel expression during Mycobacterium tuberculosis infection.* PLoS One, 2014. 9: e96427.

78. Pandey, R, et al. *MntR (Rv2788): a transcriptional regulator that controls manganese homeostasis in Mycobacterium tuberculosis.* Mol Microbiol, 2015. 98: 1168-1183.

79. Yan, S, et al. *Overexpression of Rv2788 increases mycobacterium stresses survival*. Microbiol Res, 2017. 195: 51-59.

80. Marbaniang, CN and Gowrishankar, J. *Role of ArgP (IciA) in lysine-mediated repression in Escherichia coli.* J Bacteriol, 2011. 193: 5985-5996.

81. Raghavan, S, et al. *Secreted transcription factor controls Mycobacterium tuberculosis virulence.* Nature, 2008. 454: 717-721.

82. Blasco, B, et al. *Virulence regulator EspR of Mycobacterium tuberculosis is a nucleoid-associated protein.* PLoS Pathogens, 2012. 8, e1002621.

83. Thompson, EG, et al. *Blood RNA signatures prospectively discriminate controllers from progressors early after low-dose M. tuberculosis infection of cynomolgus macaques.* J Infect Dis, 2018. Doi: 10.1093/infdis/jiy006

84. Thompson, EG, et al. *Host blood RNA signatures predict the outcome of tuberculosis treatment.* Tuberculosis, 2017. 107: 48-58.

85. Scriba, TJ, et al. *Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease.* PLoS Pathog, 2017. 13: e1006687.