A Systems Biology Approach to Identify Adjunctive Drug Targets in Two Major Mycobacterial Pathogens

by

William M. Matern

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

May, 2019

Abstract

This thesis explores novel drug targets to accelerate therapy for infections caused by the important human pathogens *Mycobacterium avium* (Mav) and *Mycobacterium tuberculosis* (Mtb). Infections with these bacterial species are notoriously difficult to treat - requiring months to years of intensive antibiotic therapy. Decreasing this length may help reduce expenditures necessary for monitoring therapy, improve patient outcomes, and reduce the significant morbidity and mortality caused by Mav and especially the global pathogen Mtb.

Bacterial antibiotic persistence has been defined as the ability of bacteria to survive in high concentrations of antibiotics without genetic mutation (ie antibiotic resistance). Thus, infections caused by Mav and Mtb are highly persistent. A major focus of this work is the discovery of mechanisms underlying this persistence phenomenon in hopes that this knowledge can be exploited to improve available therapies.

A major portion of the work is carried out using high-throughput genomic screens involving techniques such as transposon mutagenesis and transposon sequencing (Tn-seq). Statistical methods are developed and implemented to analyze this dataset with a focus on non-parametric methods. Novel discoveries include identification of the essential genes of Mav as well as particular genes that assist in bacterial survival during antibiotic exposure. Mechanisms underlying antibiotic persistence are discussed and explored in follow-up experiments guided by the high-throughput data. The mechanisms of action of antibiotics beyond well-established drug-target binding are also discussed. The results presented are relevant to the understanding of antibiotic persistence and may be informative to efforts to develop new drugs for these difficult-to-treat pathogens.

Advisors: Joel S. Bader, Petros C. Karakousis

Committee Members: Feilim Mac Gabhann, Bill Bishai

Acknowledgments

The old adage, "it takes a village to raise a child" might also be used to describe the needs of a graduate student and a thesis project. I simply could not have completed this work without the resources and advice provided by a very many people.

Firstly, I would like to thank my two PhD advisers, Joel and Petros. I am so grateful for having had the opportunity to work closely with both of you. You provided me the unique opportunity to pursue a project at the interface between two disciplines and gave me the freedom to learn by doing and go in new scientific directions. This experience has had a profound effect on my scientific thinking and made me the scientist I am today. Thank you also for your patience during the periods when my experiments weren't working and I didn't know why!

Thank you to my other two other committee members, Drs. Bill Bishai and Felim Mac Gabhann for providing independent criticism, professional support, and research guidance during my years as a graduate student.

Next, I would like thank my funding sources including The Sherrilyn and Ken Fisher Center for Environmental Infectious Diseases and the Potts Memorial Foundation. Available support for infectious disease research has recently dwindled and this work could not have been pursued without your support.

I would also like to thank the city of Baltimore and the incredibly hard-working, kind, and open community of people that I have met while living here. Baltimore has supported my scientific work during my time here in a myriad of ways. This includes the availability of quality affordable housing near my laboratory, public transportation, and the multiple centers of research and education that attract the many interesting people I have had the pleasure of meeting and working with. It is hard

not to enjoy life in Baltimore, particularly in the spring as I am writing this. I am looking forward to seeing the outcome of recent improvements in the city as Baltimore continues to attract and produce highly innovative people.

Thank you to Dr. Greg Chini and Dr. Todd Gross of the Mechanical Engineering department at the University of New Hampshire. Both of you were instrumental in my decision to pursue scientific research as a career.

Thank you to my parents and family who strongly encouraged my education from a young age.

Thank you to the members of the Karakousis lab, past and present, who trained me to work in the laboratory and directly helped with my project. Drs. Lee Klinkenberg and Dalin Rifat, you provided a huge amount of experimental assistance at the early stages of my project. Dr. Noton Dutta, you have been a continual source of advice and guidance throughout this experience and have helped keep our lab functioning. Mike Pinn, your assistance was critical during my larger experiments.

Thank you to Dr. Tom Ioerger and the Sacchetini lab at Texas A&M University for assisting with scientific interpretation, sample processing, and providing scientific advice.

The TB Center has been my home since I started in the Karakousis lab in 2014. Many past and current members have helped tremendously with my research work. To name only a few of the many: Dr. Paul Converse, Jin Lee, Dr. Shichun Lun, Dr. Korin Bullen, Dr. Laurene Cheung all provided me critical advice at various stages. A big thank you to Stef Krug who has almost single-handedly kept the upstairs (and downstairs) BSL-3 facility operational while also completing her PhD. Another big thank you to Dr. Gyanu Lamichhane for providing critical resources and scientific advice as well as the members of his laboratory for sharing their equipment and reagents and being great neighbors. The TB Center is the most open and respectful scientific community of which I am aware. And to anyone else who provided advice, a reagent, or other resource along the way that I have made the mistake of not mentioning: Thank you!

Lastly, my wife Maggie has been a continual source of motivation, inspiration, and scientific expertise that has made my work substantially better.

Contents

List of Tables

List of Figures

Abbreviations

AZI	azithromycin
cdf	cumulative distribution function
CFU	colony forming units
CLR	clarithromyin
DMSO	dimethyl sulfoxide
ecdf	empirical cumulative distribution function
EMB	ethambutol
ES	essential
FDR	false discovery rate
FWER	familer-wise error rate
GA	growth advantage
GD	growth defect
HIV	human immunodeficiency virus
INH	isoniazid
JT	Jonckheere-Terpstra
LFC	log fold change
LTBI	latent tuberculosis infection
MAC	*Mycobacterium avium* complex
Mav	*Mycobacterium avium*
MOX	moxifloxacin
Mtb	*Mycobacterium tuberculosis*
NCBI	National Center for Biotechnology Information
NE	no effect
NTM	non-tuberculous *Mycobacteria*
OD	optical density
PAS	para-aminosalicylic acid
PBS	phosphate-buffered saline

PFU	plaque forming units
RF	relative fitness
RFB	rifabutin
RMP	rifampin
SRA	sequence read archive
TB	tuberculosis
WHO	World Health Organization
WT	wildtype

Nomenclature

$U(a, b)$	uniform distribution between a and b
$X_{i,j}$	Raw read count (indexing may vary)
$\Gamma(\cdot)$	gamma function
$\Phi^{-1}(\cdot)$	inverse gaussian cdf
$\bar{q}_g$	A p-value for a particular gene g
$\psi(\cdot)$	digamma function
$\mathbf{1}[\cdot]$	indicator function
q_i	A p-value for a particular transposon site i

Chapter 1

Introduction

1.1 Tuberculosis: The global problem

The burden of tuberculosis. Despite centuries of scientific research and public health effort, tuberculosis (TB) remains a leading cause of death around the world. Mtb, the pathogen that causes TB, kills more people every year than any other single pathogen; more than HIV or malaria. The World Health Organization (WHO) estimates that roughly 10 million people developed TB in 2017 with 1.6 million dying [111].

TB and poverty. It might be argued that TB is fundamentally a disease of poverty. On a nation-state level, global TB rates are clearly associated with economic activity [53]. Furthermore, TB has been estimated to greatly reduce the global gross domestic product with approximately $65 billion removed from the global economy each year since 2015 [15]. A separate analysis estimated that TB cost to the African countries was $50.4 billion per year [57]. In the wealthy, low-burden countries a major contributor to TB is immigration from high-burden settings. In the United States roughly 66% of TB cases were among the foreign born [94]. The connection to poverty is particularly underlined by the fact that efficacious antibiotic therapy for TB has been available for decades yet millions of lives are lost every year.

High-risk groups. Other groups with high risk of developing TB include those with increased risk of exposure as well as patient with co-morbidities that make them particularly susceptible to infection. Patients with a high risk of exposure include prisoners[80] and mine-workers[92], where it is believed that poor airflow and unsanitary conditions contribute to transmission. Important

co-morbidities strongly associated with TB infection includes co-infection with HIV, which has been a major driving force in the spread of TB, particularly in sub-saharan Africa as well as diabetes mellitus, which is a particular issue in developing countries such as China and India[27].

Clinical presentation and prognosis. The typical TB patient presents with a sputum-producing cough. Infection occurring outside the lungs, generally a minority of cases[60], can have various other symptoms depending on location. Frequently the TB patient complains of night-sweats and feeling of feverishness, though objectively increased body temperature may or may not be present. In the later stages of the disease anorexia (weight loss), difficulty breathing, and hemoptysis (coughing up of blood) frequently occurs [51].

Latent tuberculosis infection (LTBI). It is well established that most patients infected with TB will not experience symptoms over their lifetime. The typical estimate for the lifetime risk of developing symptomatic, active TB given a patient is infected with the pathogen is only 10%. However, transmission of TB is widespread, which has lead to an estimated 23% of the world's population being infected with the bacterium [48]. These incredibly abundant, asymptomatic cases are classified as LTBI and are a major driver of the 10 million active cases observed each year. While efficacious (albeit intensive) therapy for LTBI has existed for decades, recent work shown that much shorter antibiotic regimens for LTBI are possible [93]. This strategy may represent an important opportunity for TB control efforts worldwide.

Goals and strategies for TB elimination. In 2015 the WHO set targets to reduce the number of deaths due to TB by 90% by 2030 and reduce the incidence by 80% as part of the Sustainable Development Goals. Additionally goals were set to eliminate financial hardship caused by utilizing health services upon contracting TB [111]. There exists at least two strategies for eliminating the mortality and economic suppression brought by TB as envisioned by the WHO. The first is to invest in general economic development in areas with high TB burden, which should eventually lead to the development of local healthcare systems and public health institutions who can effectively control TB. This, after all, was how the developed countries effectively eliminated this disease within their borders. The second strategy is to focus resources on TB control specifically, which, by their close

connections, also contributes to economic development. These two strategies highlight a critical question: How much should be invested in the development of new tools for fighting TB? The goal of this thesis to provide empirical support for a particular research direction that might, if followed through to the end, provide new TB therapies that could reduce financial and logistical burdens to healthcare delivery systems.

1.2 A brief history of TB: A therapeutic perspective

Infectious diseases and life expectancy. It is only relatively recently that infectious diseases, like TB, are *not* among the top causes of death in the developed world. In 1900 the mortality rate due to infectious diseases was 800/100,000 person-years in the United States. This declined dramatically until 1950 and since then has hovered about 50/100,000[45]. Many factors are commonly suggested to explain this accomplishment including improvements in sanitation, development of vaccines, construction of public health systems able to respond to disease outbreaks, and, most relevant for this discussion, the development of antibiotic therapy.

Discovery of the bacillus. One of the most important aspects of early infectious disease research was the isolation of pathogenic organisms. This enabled these diseases to be studied in the laboratory to much greater depth than before. By the 1870s, scientists had isolated the first organism known definitely to be the cause of a disease - the bacterium that causes anthrax. A contributor to this discovery, Robert Koch, then set out to isolate the cause of TB, for which there was evidence of an infectious cause. By 1882 Koch had isolated *Mycobacterium tuberculosis* and shown that it could cause a disease remarkably similar to TB in animals. While this discovery greatly aided diagnosis and public health measures to control the disease, it did not lead to any curative therapies until much later.

Therapy in the pre-antibiotic era. In the years before the discovery of antibiotic therapy for TB, numerous interventions were attempted to provide relief for patients. A variety of therapies were motivated by early Greek thought suggesting that the human body was composed of four "humors" (blood, phlegm, yellow bile, and black bile) and, furthermore, that disease was caused

by a loss of balance between these. Suggested therapies to rebalance the humors included dietary enrichment (such as drinking milk), the use of special plants, and - more demonstrably harmful - induced vomiting and blood-letting.

The prescriptions of rest and breathing open air was also common in the pre-antibiotic era. The sanatorium movement established refuges for TB patients to stay in bed and outdoors. Frequently the sanatoria were located in the isolated wilderness and most commonly in the mountains. In the United States the number of sanatorium beds increased to as many as 97,270 in 1942[51]. The practice of "collapse therapy" in which air is introduced into the lung cavity to disable and deflate one of the lung lobes was thought to promote healing of the lung. Variations of this technique were practiced up until the demonstration of antibiotics in the 1940s.

The beginning of antibiotic therapy for TB Paul Ehrlich, in the early 1900s had suggested that certain chemicals could be found to destroy microbes but leave the human host unscathed. By 1940 a few examples of such chemicals had been discovered - though nothing with sufficient potency to treat TB. Jorgen Lehmann, prompted by a recent report, conceived the idea of testing a derivative of salicylate against TB and eventually selected para-aminosalicylic acid (PAS)[26]. First, they showed the ability of PAS to inhibit *Mycobacterium bovis* (a cousin of Mtb). His group and other Swedish groups then showed that bacterial numbers in the sputum were greatly reduced in patients treated with PAS.

At the same time, Albert Schatz in the laboratory of Selman Waksman was testing extracts from naturally occurring soil bacteria in search of something that could inhibit the growth of Mtb on agar. He eventually discovered such an extract from the bacterium *Streptomyces griseus*, from which the drug streptomycin was isolated. Early trials demonstrated streptomycin's efficacy for treating patients.

Through the end of the 1920s and into the 1930s, Gerhard Domagk and colleagues at the Bayer Laboratories had searched for compounds capable of treating a variety of bacterial infections. Eventually, they realized that the growth of Mtb could be inhibited by certain sulfonamide compounds

with a thiazole side chain. This eventually lead them to the discovery of thioactezone. Thioacetazone was effective against Mtb but eventually found to be toxic to humans. After the war, chemists from the United States learned about thioacetazone and independently tested variations of this molecule - eventually leading to the simultaneous discovery of the highly potent and much safer isoniazid by three pharmaceutical companies - Bayer (of Germany), Hoffman-LaRoche, and Squibb Pharmaceuticals (both of the United States).

Through numerous clinical trials it was eventually discovered that the combination of these three drugs: PAS, isoniazid, and streptomycin could be combined to yield a highly potent curative therapy[35]. This "triple therapy" lasting 18 to 24 months became the standard regimen for almost 15 years[77]. However, therapy still required constant monitoring from healthcare professionals throughout the lengthy regimen, motivating additional research to find shorter regimens.

Discovery of the current regimen and subsequent attempts at further optimization. Subsequent discoveries lead to the use of the antibiotic regimen used today. Rifampin was derived from a molecule produced by *Amycolatopsis rifamycinica*, isolated from a soil sample in Italy in 1957. A series of clinical trials, many of them by British Medical Research Council (BMRC), established the usefulness of rifampin for shortening therapy[35]. Animal experiments showed that nicotinamide was active against Mtb and motivated the discovery of an analog, pyrazinamide [26]. BMRC trials established the effectiveness of pyrazinamide in shortening therapy from 9 months to 6 months. The last drug of the standard regimen, ethambutol, was shown to prevent the acquisition of drug resistance in cases where pre-existing resistance to isoniazid cannot be ruled out[51]. Therefore, while ethambutol does not reduce therapy duration, it helps to reduce the impact of antimicrobial resistance. Today, the standard therapy for curing tuberculosis is a combination of isoniazid, rifampin, pyrazinamide, and ethambutol (RHZE) for 2 months, followed by 4 months of only isoniazid and rifampin. This "short-course" therapy provides a durable cure for TB in approximately 95% of drug-susceptible cases[8].

Work since the establishment of the standard regimen has focused on finding therapies of even shorter duration. A large trial of a 4-month regimen containing moxifloxacin was found to be inferior

to the standard regimen [37]. Another large trial of a 4-month regimen containing an increased dose of rifampin is ongoing (RIFASHORT). However, large additional benefits, both financially and for reducing TB burden, could be achieved with improved drugs for TB. The White House [49], the WHO [108], and non-government organizations [34] have all made calls for the development of improved tuberculosis therapies.

1.3 Non-tuberculous Mycobacteria (NTM)

Basic Information. The non-tuberculous *Mycobacteria* (NTM) include many important human pathogens such as Mav, *M. abscessus*, *M. ulcerans*, and *M. marinum*. The NTMs are contrasted with members of the *M. tuberculosis* complex, which include *M. tuberculosis*, *M. africanum*, and *M. microti*. The *Mycobacterium avium* complex (MAC) is the most common group of NTMs to cause human disease worldwide[47]. The MAC organisms includes multiple species of NTMs with Mav being the most common in the US and perhaps worldwide[47].

Mav is an environmental pathogen that is not transmitted person-to-person[32] and is believed to be acquired directly from household water sources[29]. Patients are like infected via inhalation of aerosols containing the organism, which can be generated by common activities such as gardening and showering[44]. Mav is a robust organism and is able to survive for long periods of time without nutrients[3]. Given its ubiquity and hardiness it may be difficult to prevent patient exposure to Mav. Thus antibiotic therapy is highly important for control of this pathogen.

Disease characteristics. Mav causes disease in both immunocompromised and immunocompetent individuals. In immunocompromised persons, particularly AIDS patients, it can cause severe disseminated MAC infection. In the setting of a fully functional host response it causes at least two forms of pulmonary disease[100]. Nodular bronchiectasis (also called Lady Windermere's syndrome) is primarily observed in older Caucasian women without pre-existing pulmonary disease. Fibrocavitary disease caused by MAC is most commonly found in patients with a history of tobacco exposure and chronic obstructive pulmonary disorder (COPD) or other lung disease and affecting primarily men. It has also been noted that MAC disproportionately affects the elderly. In a study done in

Oregon, the median age of patients with pulmonary disease caused by NTMs was 68 years, with 74% of those cases caused by MAC[18].

The typical patient infected with Mav in the lungs may present with various symptoms. A long-lasting cough is commonly present. Fatigue, shortness of breath, coughing up blood, night sweats, and anorexia can also all be present. In the case of immunocompromised patients (such as by HIV/AIDS) bacteria can be present in the blood (bacteremia). Symptoms can include fever, night sweats, chronic diarrhea, weakness, and abdominal pain.

Epidemiology. Unlike the plethora of data for TB, there are no available statistics for the global incidence of Mav infection. Incidence and mortality are particularly difficult to estimate as infection with NTMs, including Mav, are generally not reportable to public health authorities [82]. Additionally, in most datasets Mav is rarely distinguished from other members of the MAC. Some statistics are available for particular countries. In the US, prevalence of NTMs has been estimated as 27.9/100000 [91]. MAC (not specifically Mav) may account for 80% of these infections [56]. Furthermore, Mav is believed to be the dominant disease causing member of the MAC pathogens in North America [47]. Based on the available data, it can be reasonably inferred that Mav is likely the most common NTM to cause disease in the U.S..

Treatment First-line therapy for MAC infections is recommended to include clarithromyin (CLR) or azithromycin (AZI), rifabutin (RFB), and ethambutol (EMB). The recommended duration of antibiotic therapy for immunocompetent individuals is at least 12 months after bacteria are no longer present in sputum[41]. The treatment success rate for MAC infection in the immunocompetent host, as measured by culture negativity within a year of starting therapy, is approximately 82% for the standard regimen [42]. The antibiotic therapies available for MAC have significant side effects including nausea, vomiting, and skin rashes leading to reductions in patient doses[90]. Achieving shorter duration and more efficacious therapy with lower doses is essential for reducing drug intolerance and improving patient outcomes.

1.4 Why are current drugs so ineffective at treating *Mycobacteria*?

The duration of antibiotic therapy for Mav and Mtb infections is significantly longer than that of other bacterial infections. For example, in the absence of antibiotic resistance, community acquired pneumonias (non-mycobacterial) are generally treated with 5-7 days of antibiotics [72] compared to the months to years of therapy required for these mycobacterial infections.

The proposed theories to explain the long therapy duration invoke the structure of the lesions observed during tuberculosis. The pathology of pulmonary TB includes the formation of granulomas in the lungs. These granulomas can vary in appearance and properties. Frequently granulomas, in the absence of therapy, will form necrosis at the center due to dying host cells. Furthermore, granulomas can develop fibrosis at their outer edges. Given the organized and non-vascularized structure of granulomas there may be reduced availability of nutrients and (in the context of therapy) antibiotics within these lesions relative to the blood stream and surrounding tissue.

Two popular theories exist to explain the exceptionally lengthy therapy required for Mtb infections. One hypothesis, here called the "persister hypothesis", postulates the existence, within the host, of a special population of bacteria refractory to current antibiotics [115]. As discussed later, reduced nutrient availability within granulomas may contribute to the development of these special bacteria. Another hypothesis, here called the "drug-permeability hypothesis", proposes that currently used drugs do not efficiently penetrate into the closed lesions believed to contain the bacteria [22]. This hypothesis suggests that efforts to accelerate therapy should focus on increasing penetration of drugs through changes in chemistry or drug delivery and that new drug development should focus on permeating areas where current drugs do not. There is evidence supporting both theories of persistence but no clear demonstration of which types of interventions have stronger effects on treatment duration in people. However, given the substantial evidence supporting the existence of persisters *in vivo* and the observed reduction in therapy duration when a persister-targeting drug was added to standard therapy in an animal model [43], we will focus here on exploring interventions to target persister bacteria.

1.4.1 Bacterial persister hypothesis

Substantial work on exploring persister bacteria has been performed with the model organisms *E. coli*. In *E. coli*, a previous study provided evidence that two different persister states can be defined *in vitro* [5]. Type I persisters, based on the behavior of the *hipA7* mutant, are non-growing bacteria produced during stationary phase. In general, it might be said that Type I persisters are induced by the environment, and particularly by reduced nutrient availability. Type II persisters, based on the *hipQ* strain, are slow-growing and generated rarely from non-persisters during exponential growth. Wakamoto and colleagues showed direct evidence for the existence of a population of cells matching the Type II persister definition in wildtype *Mycobacterium smegmatis*[106]. This suggests that one factor that may allow *Mycobacteria* to survive during antibiotic exposure in a patient is the stochastic generation of a special population of slow growing bacteria.

There is also a large body of work suggesting that various stress conditions can induce an antibiotic tolerant state in *Mycobacteria* which suggests the existence of Type I persisters. There is strong evidence that Mtb-containing lesions (granulomas) are hypoxic [104] and hypoxic bacteria have been shown to be significantly more tolerant of multiple antibiotics including the first-line TB drugs isoniazid (INH) and rifampin (RMP) [36]. Nutrient starvation (immersion in Phosphate-Buffered Saline) has also been shown to induce tolerance in Mtb to many antibiotics including INH and RMP [36, 112]. It has been speculated that nutrient starvation may occur in the fibrotic areas of granulomas though it is unclear which nutrient(s) may be limiting[68]. Mav was shown to be significantly more tolerant of both isoniazid and pyrazinamide in another *in vitro* stress model [4]. Additionally, bacteria were shown to be more tolerant of CLR and EMB in specialized culture conditions to produce a biofilm [75]. However, little evidence exists that biofilms states occur in vivo. Despite the clear demonstration that the environment and antibiotic susceptibility are closely linked, the mechanisms and pathways underlying this environmentally-induced persistence, especially in *Mycobacteria*, remain to be worked out. Knowing these mechanisms could provide a useful guide in designing shorter regimens to treat infections with Mav and Mtb.

1.5 Experimental opportunities: High-throughput biology

High-throughput biotechnologies can help to identify candidate antibiotic persistence mechanisms for further exploration. In this work we have a particular focus on using high-throughput genomics technologies (including transposon mutagenesis) to identify genes involved in the antibiotic persister phenotype. This was enabled by recent developments in DNA sequencing which can provide the investigator the base-by-base sequence of input DNA molecules. This work utilizes both short read technology primarily from Illumina (eg. HiSeq, MiSeq, NextSeq) and long read sequencing technology from Pacific Biosciences (eg. RSII).

In analyzing the large amounts of data produced from a modern DNA sequencer a number of important considerations must be made. First is the consideration of how one should normalize the data. The efficiency by which a sequencer produces reads inherently varies between samples. This can be caused by many factors including differences in sample quality and the length of DNA molecules input to the machine. The most common method is normalization to the total number of reads. The resulting fractions can then be compared between samples. However, other methods for normalization exist such as division by read count from a DNA species with a known stable value (e.g. a "housekeeping" gene).

An additional consideration for the interpretation of sequencing data is the method employed to compare two sample groups. Typically the chosen method consists of some set of mathematical operations achieving some statistical guarantees under certain assumptions. The method employed can be as simple as a t-test for comparing the (normalized) levels of a gene between two samples. Alternatively, sophisticated methods employing more obscure distributions such as the negative binomial can be employed. This thesis will employ two styles of method for analysis. Chapter 2 will discuss the use of a modified negative binomial distribution employed to model read counts sampled from a highly non-Gaussian distribution. Chapters 2-4 will also discuss model-free (or non-parametric) methods which make no assumptions about the particular form of the underlying distribution. These non-parametric methods can provide substantially more robust results as fewer assumptions need to be made about the input data.

High-throughput biology, while enabling large numbers of different molecules to be measured at the same time, has also opened the door for large amounts of noise to permeate the results. The sources of noise are not unique to high-throughput data but, given the large numbers of different tests involved, can be difficult to manage. While statistical methods can help to reduce the impact of this noise on the results, no statistical method can fully eliminate it. Therefore any result, no matter how significant, must be followed up with additional experimentation, preferably with an entirely unrelated experimental methodology. Given the large number of results from high-throughput techniques, they could not all be independently tested with low-throughput techniques. However, some of the most important outcomes were validated and support the conclusions gleaned from the high-throughput data.

1.6 Outline of Thesis

Chapter 2 of this thesis will discuss work to provide genomic-scale data for a strain of Mav used in many prior laboratory studies. Of particular note, complete genome sequence of this strain as well as computer annotations for the likely locations of the genes and other features in the DNA are provided. A list of genomic features most important for this organism to achieve growth in the laboratory is also provided. These data may be most useful to drug developers who wish to find new ways to treat this organism in the human host as well as for laboratory scientists who may be considering the disruption of certain features for further study. Chapter 3 will concern the employment of high-throughput techniques to identify elements of this organism that play a role in the efficacy of antibiotics. As previously discussed, antibiotic persistence (ie the survival of bacteria in the presence of antibiotics) remains mysterious. This work attempts to shed light on the mechanisms through which persistence is achieved by identifying the genetic features with roles to play. These data may be of particular interest to those looking to optimize antibiotic regimens for the treatment of patients. Chapter 4 discusses work to identify genomic features in the global pathogen Mtb that attenuate the effect of antibiotics. The role that environmental conditions can play on the organism's ability to survive under antibiotic insult is also explored. Furthermore, it

is shown that a particular gene can boost the effectiveness of one of the first-line drugs used to these infections with this organism. Chapter 5 discusses the results from the previous chapters and provides some conclusions and directions for future work.

Chapter 2

Identifying the essential genes of *Mycobacterium avium*

2.1 Objectives

The availability of fast and inexpensive DNA sequencing has enabled the genomes of many organisms to be determined. A full genome, without gaps or missing sequences is a useful guide for many follow-on experiments. This includes cloning of genes for investigation of function and for interpretation of other sequencing data such as RNA-seq or transposon sequencing (Tn-seq). Determination of the full genome for an organism can now be considered one of the first steps in characterizing the organism.

Tn-seq is a methodology to quickly characterize gene function in a highly efficient parallelized way. Most relevant for this chapter is the use of Tn-seq to identify genes required for growth of the bacterium. These genes required for growth or "essential genes" - as they are commonly called - are of particular interest for the identification of candidate drug targets and for the further characterization of the genome. For example, a pharmaceutical company searching for anti-microbials may wish to focus on identifying inhibitors for essential genes. Additionally, essential genes are not possible to study through the construction of knockout strains, as is commonly done for non-essential genes. Thus, the identification of essential genes can be helpful in the choice of experimental method.

While extensive efforts have been made to identify the essential genes of several other Mycobacterial species, no complete list of the essential genes in Mav currently exists. An attempt was made in subspecies *paratuberculosis* (a subspecies that infects cattle) but this did not use a sufficiently

diverse transposon mutant pool to assess the essentiality of individual genes. It was anticipated that this knowledge, using a human-derived strain, would be very helpful for guiding future efforts to characterize this important pathogen.

Therefore, we have generated the complete genome sequence of a patient-derived strain of Mav which is commonly used in preclinical work. This allows integration with the significant knowledge that has already been generated with this strain. We have completed this sequence without using a reference genome (ie *de novo* assembly) partly due its accessibility using modern techniques as well as the significant anticipated gain in quality for the final sequence given the genomic diversity of Mav. Furthermore, we have generated transposon mutant pools of sufficient diversity to provide high quality identifications of the essential genes of Mav. This has enabled us to provide a high-quality list of the essential genes in Mav.

2.2 *De novo* assembly and annotation of a novel Mav genome

Substantial portions of this section appeared recently in the journal Scientific Data.[74]

Mav strain 109 (MAC109) was isolated from the blood of an AIDS patient and has been used frequently in preclinical investigations [33, 65, 13, 17, 46]. In our efforts to study the essential genes of Mav via transposon mutagenesis (using ϕMycomarT7 [87]), we observed the highest insertion efficiency with this strain of those we tested (see section 2.3). To simplify the follow-on analysis as well as enhance the potential utility of MAC109 for future studies focused on Mav genetics, including RNA-seq and ChIP-seq analyses, we decided to sequence the MAC109 genome. Given the substantial genetic heterogeneity observed between M. avium isolates[99] (such as the frequent inclusion of plasmids), we performed *de novo* assembly. We also provide Gene Ontology (GO) annotations, which have proven useful in exploratory analysis of the roles of individual genes.

2.2.1 Experimental and Computational Methodology

This section discusses the experimental and computational methodologies used to assemble the genome of MAC109.

Bacterial strain. MAC109 was received on agar as a gift from Dr. Luiz Bermudez (Oregon State University). A single colony was inoculated into Middlebrook 7H9 broth supplemented with 10% OADC, from which stocks were made. These stocks were frozen and subsequently passaged into fresh broth for generation of bacterial samples for DNA extraction.

DNA Extraction for short read sequencing. DNA was extracted from MAC109 using two different methods. For Illumina sequencing we used a fast DNA extraction protocol that produces fragmented DNA. First, a 2 mL O-ring tube was filled with 1 gram of 0.1 mm zirconia/silica beads (Biospec #11079101z). Then 0.6 mL of liquid containing bacteria was transferred to each tube of beads. 2uL of RNase A (10 mg/mL, Thermo-Fisher #EN0531) was then added to each tube. Tubes were bead-beaten for 1 min at max speed. Tubes were then incubated at room temperature for 10 min to degrade RNA. 0.6 mL of saturated phenol solution was then added to tubes. Tubes were then briefly vortexed and spun for 2 min at 16,000g. The aqueous phase was transferred to clean microcentrifuge tubes and an equal volume of chloroform/isoamyl (24:1, v/v) was added. Tubes were briefly vortexed and then centrifuged for 2 minutes at 16,000g. 400 uL of the aqueous phase was then transferred to clean microcentrifuge tubes. 40 uL of 3 M sodium acetate was added and tubes were inverted to mix. 880uL of 100% ethanol was added and tubes were inverted to mix. If the DNA pellet was visible after mixing then tubes were spun at 16,000 g for 2 min, otherwise they were spun for 15 min. Supernatant was removed and 700 uL of 70% ethanol was added to wash the pellet. Tubes were centrifuged for an additional 30 s. Supernatant was removed and pellet was allowed to dry at room temperature for 15 min. Pellet was then dissolved in 100 uL of TE buffer. A nanodrop was used to determine concentration.

DNA Extraction for long read sequencing. For PacBio sequencing, which requires high molecular weight DNA, bead beating could not be used and a high molecular weight DNA extraction protocol for mycobacteria[7] was adapted for this purpose. First, a bacterial pellet was obtained in a 2 mL Eppendorf tube in 500 uL TE buffer (pH 8.0). We then added an equal volume of chloroform/methanol (2:1, v/v) and rocked on a platform rocker for 5 min. Then we centrifuged the suspension at 16,000 g for 2 min. The aqueous and organic phases were removed by pipetting, leaving

only the band of bacteria at the interface. The delipidated bacteria were then placed in a dry bath at 55 °C for 10–15 min to remove traces of the organic phase. 550 uL of 10 mg/mL Lysozyme (Sigma-Aldrich, dissolved in TE buffer) was then added to the bacteria. The pellet was gently resuspended with a pipette. Tubes were then placed in a 37 °C incubator overnight. On the following day 120 uL of 5% SDS (w/v) and 8uL proteinase K (10 mg/mL, Thermo-Fisher #EO0491) were added to the tubes. Tubes were then mixed by gentle inversion and incubated at 50 °C for 3 hours in a dry bath. An equal volume (750 uL) of PCI (phenol solution/chloroform/isoamylalcohol, 25:24:1, v/v) was then added and gently inverted on a rocker plate for 30 min. Tubes were then centrifuged at 16,000 g for 5 min. ∼600 uL of the aqueous phase was then transferred to a clean microcentrifuge tube. 2 uL of RNase A (10 mg/mL, Thermo-Fisher #EN0531) was added and tubes were inverted to mix. Tubes were then incubated at room temperature for 10 min to remove contaminating RNA. Chloroform/isoamylalcohol (24:1, v/v) was then added and tubes were inverted for 30 s to mix thoroughly. Tubes were spun for 2 min at 16,000 g and 400 uL of the aqueous phase was transferred to a clean microcentrifuge tube. 40 uL of 3M NaCl was added and tubes were mixed by inversion. 0.8 mL of 100% EtOH was added and tubes were mixed by inversion. If DNA pellet was visible by eye then tubes were spun at 16,000 g for 2 min at room temperature, otherwise they were spun for 15 min. The supernatant was then removed from the pellet and 0.8 mL of 70% ethanol was added to wash the pellet. Tubes were then spun at 16,000 g for 5 min and supernatant discarded. DNA pellets were dried for 15 min at room temperature. After drying, pellets were resuspended in 100 uL TE buffer.

Sequencing and *de novo* assembly of MAC109 genome. Library preparation and sequencing were performed by the Deep Sequencing and Microarray Core at the Johns Hopkins School of Medicine. Library preparation was done with the Illumina TruSeq DNA Nano kit with target insert size of 350 bp. Short reads were generated using an Illumina Miseq (2×75 bp). After clipping adapters (internally by Illumina software), there were 3,117,540 reads, which varied in size from 35 to 76 bp (472,180,593 total bases sequenced). Long reads were generated with a PacBio RSII after library preparation with the SMRTbell Template Preparation kit 1.0 and target read size of 10-20

kb. Raw PacBio data (*.h5) were converted to subreads (fastq format) with SMRT Tools v5.1.0 (bax2bam v0.0.8, bam2fastq v1.1.1). Raw reads from the Illumina Miseq and Pacbio RSII machines are available in sequence read archive (SRA).

Reads smaller than 300 bp were filtered out using Trimmomatic v0.36 [11]. After filtering, there were 151,792 subreads and 1,365,362,111 total bases sequenced. Assembly of the genome was performed using Unicycler pipeline v0.4.4 in conservative mode [110]. Dependencies for Unicycler were satisfied with SPAdes v3.11.1 [6], racon v1.2.1 [103], bowtie2 v2.3.4.1[62], and pilon v1.22 [107]. Unicycler was run in hybrid mode, allowing the use of both the Miseq (short, accurate) reads and PacBio (long, less accurate) reads. This produced an assembly with four circular contigs. However, one of these contigs was nearly identical to the genome of the bacteriophage PhiX174. It is highly unlikely that this phage is part of the MAC109 genome and much more likely that this contig was assembled from contaminating reads from the PhiX library run alongside our sample on the Miseq. The PhiX contig was therefore removed from the Unicycler assembly. Additionally, preliminary annotation revealed that the sequence of one of the plasmids started inside a gene. Since this complicates downstream processing, the start of the sequence was moved accordingly. Bandage v0.8.1 [109] was used to visualize the final genome assembly (Fig. 2.1a).

Gene and Gene Ontology (GO) Annotations. Gene annotation was done with the Prokaryotic Genome Annotation Pipeline (PGAP) available from the National Center for Biotechnology Information [96]. Proteins were given GO annotations using the PANNZER2 webservice with default settings [98]. We submitted our protein sequences for GO annotation on 2018/06/19 (the PANNZER2 databases are updated monthly).

Genome Comparison. The nucmer program from MUMmer v4.0.0beta2[73] was used to compute SNPs and dot plots between the MAC104, TH135, and MAC109 genomes. Settings for the full genome dot plots (Fig. 2.1b and c) were "nucmer –maxmatch -l 20" followed by "delta-filter -l 2500 -m". For dot plots comparing the TH135 and MAC109 plasmids, settings were "nucmer –maxmatch -l 10" followed by "delta-filter -l 1000 -m". Plots were made with gnuplot v5.2 patchlevel 4 (Fig. 2.1b–d). To confirm the subspecies of MAC109 we also used nucmer (with settings: –maxmatch and

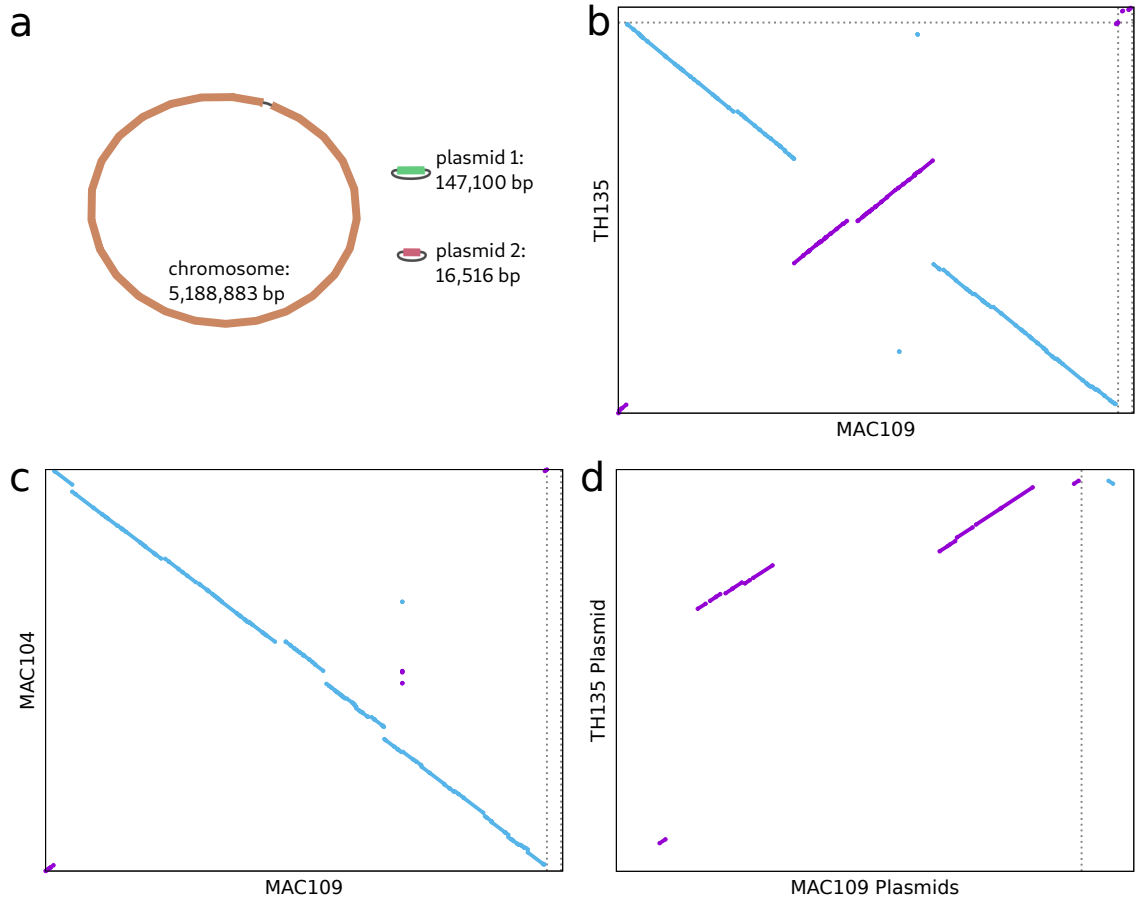

Figure 2.1: Summary of MAC109 genome assembly and homology. (a) MAC109 genome assembly containing 3 complete circular contigs. (b,c) Dot plots comparing TH135, MAC104, and MAC109 genomes assemblies. Dotted lines separate the replicons of each strain (TH135 has a single plasmid, MAC104 lacks plasmids). (d) Dot plot comparing the plasmid from the TH135 genome and the two plasmids from the MAC109 genome.

–minmatch = 10) to check for the presence of insertion elements IS900, IS901, IS1311, and DT1.

2.2.2 Results

The genome of MAC109 consists of a circular chromosome of size 5,188,883 bp and two plasmids of sizes 147,100 bp and 16,516 bp (Fig. 2.1a), and approximate multiplicities (estimated by Unicycler) of 1.76x, 4.92x, respectively. We have named the larger plasmid pMAC109a and the smaller plasmid pMAC109b. Based on the presence of insertion element IS1311 (GenBank accession no. U16276) and absence of IS900 (accession no. X16293), IS901 (accession no. X59272), and DT1 (accession no. L04543) we confirm that MAC109 belongs to the Mav subspecies *hominissuis* [89]. PGAP identified 4,841 protein coding sequences, 53 RNA genes (including 46 tRNAs, 4 rRNAs, and 3 ncRNAs),

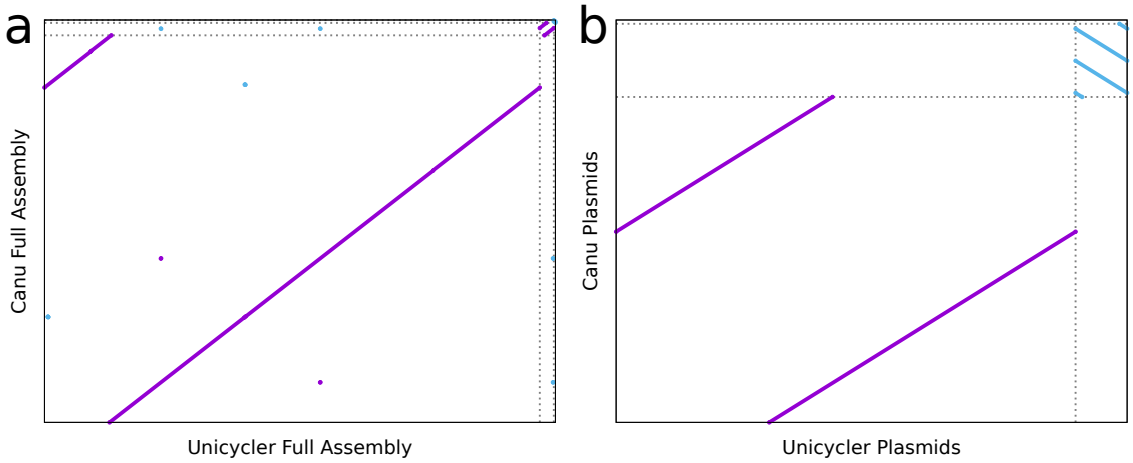

Figure 2.2: Comparison of genome assembly of MAC109 output by Canu vs Unicycler. (a) Dot plot comparing full Unicycler (3 contigs) assembly vs full Canu assembly (4 contigs). Dotted lines separate the replicons of each strain. (b) Same comparison but with the largest contig from each assembly removed. This provides a higher resolution comparison of the small contigs.

and 191 pseudogenes. The annotated genome and raw reads are available under Genbank accession numbers CP029332-CP029334.

We compared the MAC109 genome to those of Mav strains TH135 and MAC104. MUMmer/nucmer estimated 32,974 SNPs relative to MAC104 and 56,751 SNPs relative to TH135 (there were 46,685 SNPs between TH135 and MAC104). Dot plots of these comparisons are presented in Fig. 2.1b–d. Notably, a number of large-scale inversion have occurred in these strains. Additionally, the pMAC109a plasmid from the MAC109 assembly shares significant regions of similarity with the plasmid in TH135, although there are large distinct regions as well (Fig. 2.1d). PANNZER2 provided a total of 17,292 GO annotations for 3860 unique proteins (among a total of 4841 proteins). The full list and descriptions of the GO terms can be found at `http://www.geneontology.org/page/download-ontology`.

2.2.3 Validation of assembled genome

To provide support for our assembled MAC109 genome we show that using an entirely different assembler (Canu v1.7.1[59]) yields an almost identical genome thus supporting our reported genome. However, unlike Unicycler, Canu is not a hybrid assembler and does not attempt to circularize contigs. Therefore, a few minor differences are expected. In particular, the assembled contigs are

linear and might have identical regions at the ends of contigs when the underlying DNA molecules are circular. Also, Canu does not attempt to move the origin of circular contigs after assembly. Therefore, the Canu-derived contigs are likely to have different origins than the Unicycler contigs. After running ("canu genomeSize = 5.35 m -pacbio-raw pb_reads.fastq.gz", other settings set to defaults), Canu assembled 4 linear contigs (Lengths: 5,207,511 bp, 167,345 bp, 37,619 bp, and 1974bp). To compare the Canu assembly with the Unicycler assembly we used MUMmer ("nucmer –maxmatch -l 20" followed by delta-filter to filter out small matches). Fig. 2.2a shows a dotplot comparing the entire genomes ("delta-filter -l 2500"). Figure 2.2b compares the smaller contigs (plasmids) of the Unicyler and Canu assemblies ("delta-filter -l 1000"). This shows that the Canu contigs repeat themselves at the ends (and in the case of contig 3, repeats occur multiple times), as expected. Secondly, it can be seen that contigs 1-3 are nearly identical to the Unicycler contigs, supporting the accuracy of our assembly. However, Canu assembled one additional contig (Canu contig 4, length = 1974bp) relative to Unicycler. Canu contig 4 was noted to have a high TA content, although *Mycobacteria* are known to be GC-rich. To test whether the Canu contigs are actually present in the MAC109 genome, we mapped our collected Illumina reads, which were not used in the assembly by Canu, to estimate the multiplicity of each contig. We used bowtie2 v2.3.4.1 with default settings. Overall, >97.9% of the Illumina reads mapped to the Canu contigs. We then used samtools v1.5 [64] ("samtools sort -o file2.bam file1.sam" followed by "samtools depth file2.bam") to calculate the depth at each position and averaged the depth across each contig to calculate the coverage using a short python3 script. Coverage was 83x for contig 1, 128x for contig 2, and 177x for contig 3. 0 reads mapped to the 1974bp contig output by Canu. Therefore, contig 4 appears to be an entirely erroneous contig. Therefore, we have shown that our assembled genome is robust to changes in the assembler, although Canu produces an erroneous contig. This supports the quality of our genome sequence for MAC109.

2.3 Construction of genome-wide mutant pools

Transposon sequencing (eg TraDIS[63], Tn-seq[24], INseq[39]) has been extensively used to profile haploid genomes and identify gene disruptions that affect growth under various conditions. Of potential interest in drug development are those drug targets which profoundly disrupt growth (ie "essential" genes). To inform future research and Mav drug development efforts we have identified genes critical for bacterial growth *in vitro*. To enable this, we developed a statistical approach for calling genes, based on robust techniques, which is designed to be less susceptible to common sources of error than existing techniques. We report our predictions of the essential genes of Mav and compare with predicted essential genes in a cousin, Mtb.

2.3.1 Experimental Methodology

Strains. Mav strains MAC109, MAC104, OSU3388 were a gift from Dr. Luiz Bermudez (Oregon State University). MAC101 (Chester, ATCC 700898) was a gift from Dr. Eric Nuermberger (Johns Hopkins School of Medicine). Individual colonies of each strain were isolated and regrown to make stocks used in the described experiments. MAC101 was seen to form both translucent and opaque colonies. Both an opaque (MAC101o) and a translucent (MAC101t) colony were isolated and used for stocks. ϕmycomarT7 was propagated and titered as previously described[71]. Final titers used for transformations exceeded 10^{11} plaque forming units (PFU)s/mL.

Media and buffers. To make 7H11 agar 10.25 grams of 7H11 w/o Malachite Green powder (HiMedia Cat No. 511A) was added to 450mL deionized water. 5mL 50% glycerol was then added before autoclaving. Hot agar was cooled to 55°C before addition of 50mL OADC enrichment and 1.25mL 20% Tween-80. To make 7H9/10% OADC: 2.35g 7H9 powder was added to 450mL deionized water. After sterilization (via autoclaving at 121°C or by passing through a 0.22um filter) 50mL of OADC enrichment (Becton-Dickinson) was added. Unless otherwise specified, no Tween-80 or glycerol was included. To make 7H9/50% OADC: Recipe identical to 7H9/10% OADC but using 250mL water and 250mL OADC. To make PBS-Tw: 1.25mL filter-sterilized 20% Tween-80 was added to 500mL sterile phosphate-buffered saline (PBS). To make MP Buffer: final concentrations

are 50mM Tris-HCl (pH 7.5), 150mM NaCl, 10mM MgSO4, 2mM CaCl2. Autoclave individual components before combining.

Testing of transformation efficiency of Mav strains. Five strains of Mav (MAC109, MAC104, OSU3388, MAC101o, MAC101t) were tested for transformation by ϕmycomarT7. For transformation, strains were grown in 150mL of 7H9/10% OADC. After OD of each strain reached 0.32 – 0.89, 100mL of cultures were equally split into two 50mL conical tubes. Bacteria were pelleted via centrifugation and reimmersed in 10mL MP buffer. Bacteria were pelleted again via centrifugation and reimmersed in 4.5mL MP Buffer. 0.5mL of MP Buffer (negative control) or ϕmycomarT7 stock (approximately 10:1, phage:bacteria) was added to each tube. Tubes were incubated for two days shaking at 37°C. Bacteria were then pelleted via centrifugation and reimmersed in PBS-Tw. Bacteria were spun down again and reimmersed in 1mL of PBS-Tw. Transformed bacteria and negative control for each strain were then diluted in PBS-Tw and plated on 7H11 with and without 50ug/mL kanamycin for determining transformation efficiency and background resistance.

Generation of transposon mutant libraries in MAC109 In preliminary experiments we found that MAC109 growth increased in broth containing higher concentrations of OADC. We suspect the oleic acid in OADC is the key to achieving this, based on previous reports[28]. 5 independent transposon mutant pools were generated. MAC109 was grown in 700mL 7H9/50%OADC to OD 2.1 in two 1.5L roller bottles shaking at 37°C. Based on previous results (data not shown) we estimated the initial bacterial density based on optical density to be 4×10^8 CFUs/mL (used for calculating volume of phage stocks to add). Bacteria were aliquoted to 12-50mL conical tubes and centrifuged (2000g for 5 minutes) and supernatant removed. 5mL MP Buffer was added to each tube and bacterial pellet was reimmersed. Pairs of tubes were pooled yielding 6 10mL aliquots. Samples were then centrifuged (2000g for 5 minutes) and supernatant removed. Phage (10:1, phage:bacteria) was then added to all tubes except no-vector control. MP Buffer was added to all tubes to a final volume of 5mL and bacterial pellets were dispersed via pipette. Bacterial/phage mixtures were then placed on a shaker incubator (37°C) for two days. Tubes were then centrifuged (2000g for 5 minutes) and supernatant removed. 10mL PBS-Tw was then added and the bacterial pellet was dispersed via

Table 2.1: Transformation efficiency of various strains of Mav with ϕmycomarT7.

Strain	Transformation Efficiency	Background Resistance
MAC101t	1.0E-05	1.5E-07
MAC101o	6.3E-08	1.4E-08
MAC104	3.7E-07	2.9E-08
MAC109	1.5E-05	1.5E-07
OSU3388	1.9E-07	1.4E-07

pipette. Tubes were then spun down again (2000g for 5 minutes), supernatant removed, and 1 mL of PBS-Tw was used to reimmerse pellets.

50uL of each tube of washed transformants (or no-vector control) were diluted and plated on 7H11 plates, with or without 50ug/mL kanamycin, to determine transformation efficiency and background resistance. The remainder of the cultures were plated on 7H11 containing 50ug/mL kanamycin in Pyrex baking dishes (15" x 10", 500mL agar per dish, 1 tube per dish). After 7-10 days colonies were scraped from each dish and dispersed in fresh 7H9 broth and frozen in aliquots at -80°C for later use.

2.3.2 Results

To identify a suitable strain of Mav for genome-wide mutagenesis we evaluated the ability of ϕmycomarT7 to transform common laboratory strains. Transformation efficiency and spontaneous resistance rate (background) were estimated via CFU counts and are provided in Table 2.1. MAC109 was observed to have the highest transformation efficiency with only ∼1% background. Therefore, we decided to proceed with transposon mutagenesis using this strain. Upon transformation, we estimated each of our five independent MAC109 transposon mutant libraries contained between $2.2\breve{}4.4 \times 10^5$ unique insertion events for a combined total of 1.2×10^6 unique events with ∼2% background.

2.4 Identifying essential genes

2.4.1 Collection of raw data and processing

DNA was extracted from one aliquot of each transposon mutant pool as generated in section 2.3 using a previously described gDNA extraction protocol for short read sequencing 2.2. We adapted

Figure 2.3: Tn-seq library preparation produces DNA molecules for sequencing as shown here. The adapters at each end are involved in attachment to the flow cell. The ends of the transposon (both 3- and 5-prime contribute) are adjacent to a gDNA segment which allows for the location of the transposon insertion within the genome to be determined. Lastly a UMI is used to identify (and remove) PCR duplicates originating from the same bacterium - thus reducing PCR bias.

a previously published library prep protocol[69] to prepare libraries for sequencing. Adaptations include the use of magnetic beads for purification and library size selection as well as changes to PCR conditions (for details see appendix B). This produces DNA sequencing libraries as shown in figure 2.3. Libraries were sequenced (2×75bp) on an Illumina HiSeq 2500 by the Johns Hopkins GRCF High Throughput Sequencing Center. 5 independent libraries were sequenced yielding between 2,194,085 – 4,381,545 reads per library for a total of 18,197,728 paired-end reads. The raw reads are provided in SRA under accession number: PRJNA527645.

As described above, the MAC109 genome contains two plasmids in addition to the bacterial chromosome. We adapted the TRANSIT pre-processor (tpp)[25] to allow for mapping to multiple contigs. Our changes were included in the release of TRANSIT/tpp v2.4.0. We used tpp v2.4.0 to map all reads to the MAC109 genome. Command for processing raw reads: tpp -himar1 -bwa -bwa-alg aln -ref MAC109.gb -replicon-ids "CP029332,CP029333,CP029334" -reads1 TnPool_1.fastq -reads2 TnPool_2.fastq -window-size 6 -primer AACCTGTTA -mismatches 2. After PCR duplicate removal, a total of 10,597,261 unique reads mapped to the genome and were used for analysis.

Confirmation of low permissibility sites. It was previously shown that the Himar1 transposon/transposase system has a reduced rate of insertion in sites with the sequence motif [CG]GNTANC[CG][24]. Indeed, our results confirm that insertion into these low permissibility sites is much less likely than other sites (Figure 2.4). While we are close to hitting all possible insertion sites in the genome not matching this motif (ie achieving saturation), a substantial fraction of the low permissibility sites in the chromosome are unoccupied in all five libraries. This effect is less apparent in the plasmids likely due to their multiple copy number.

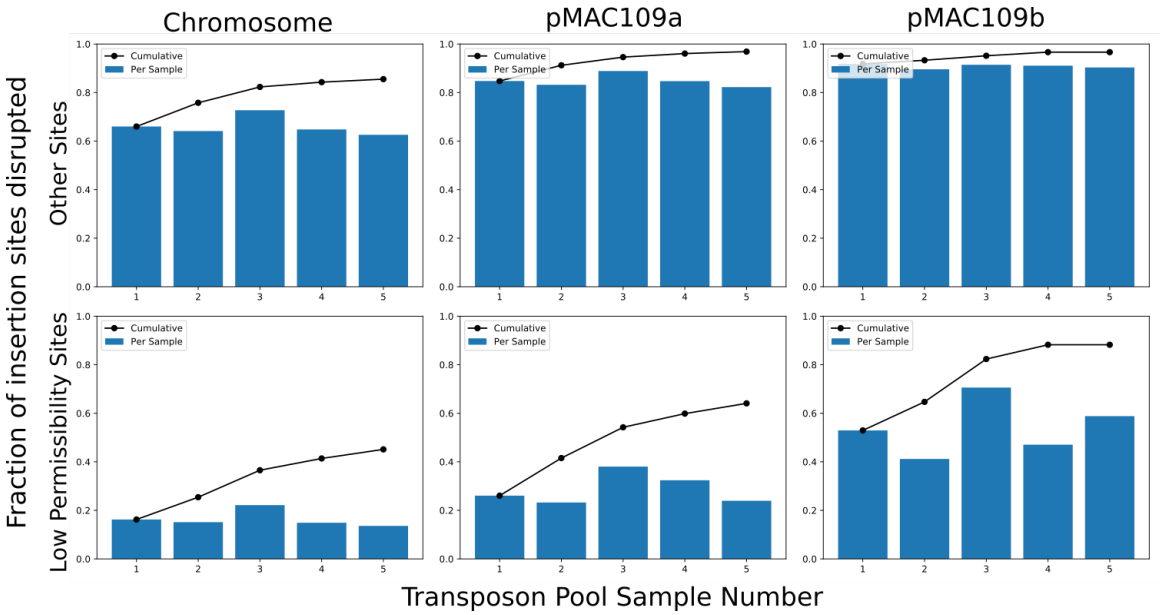

Figure 2.4: Each barplot shows the fraction of potential Himar1 insertion sites (TA dinucleotide) observed to have sustained at least one insertion in each independent pool of mutants for each replicon of the MAC109 genome. The line plots indicate the cumulative fraction of occupied insertion sites. Notably, the fraction of unique sites occupied saturates for sites not matching the previously defined sequence motif for low permissibility sites ([CG]GNTANC[CG]). However, sites matching this motif can be seen to be near saturation only in the case of the small plasmid (pMAC109b).

2.4.2 Statistical analysis

Overview. We used a previously suggested labelling scheme[23] to annotate each gene of MAC109. A gene is labelled no effect (NE) if a transposon insertion in any of its potential insertion sites causes no effect on growth. A gene is labelled growth defect (GD) if it contains at least one insertion site such that upon transposon insertion it results in a decrease in bacterial growth. A gene is labelled growth advantage (GA) if it contains at least one insertion site that *increases* bacterial growth. A gene is labelled essential (ES) if it contains at least one insertion site that results in a 20-fold loss in viability.

To classify the features of the MAC109 genome with the above schema we have designed a robust procedure. At a conceptual level, our analysis pipeline proceeds in two steps. First, insertion sites with no effect on growth (NE) are approximately identified with a rank-based filter procedure. Second, the counts at the insertion sites identified by the filter are assumed to approximate the null

distribution and used for statistical hypothesis testing. For identification of ES genes, the approximate null distribution is fit to a zero-inflated negative binomial distribution, which is then scaled and used for hypothesis testing. For identifying the GD and GA sites, the empirical cumulative distribution function (ecdf) is used for hypothesis testing. Stouffer's method is used to combine p-values from multiple replicates and multiple sites. Lastly, multiple hypothesis correction is performed.

Rank-based filter procedure. Here we will briefly describe the key assumptions and calculations supporting our approach. We assume that all mutants with a transposon insertion at the same site will have identical growth rates (ie the growth rate is entirely defined by the insertion site). This implies that all insertion sites with no effect on growth will be identically distributed. We also assume that not more than 40% of insertion mutants have a growth defect and not more than 15% of mutants have a growth advantage (and therefore at least 45% of mutants will have a growth rate that is identical to wildtype). See figure 2.5 for a conceptual diagram. We have chosen these thresholds based on previous predictions in Mtb[24] suggesting that disruption of 15% of genes (a little less than 15% of insertion sites) cause a growth defect and 8% cause a growth advantage. We have added a large margin of error to ensure conservatism. Note that if some of the identities of insertions mutants with growth rates identical to wildtype were known ahead of time we could simply use the distribution of the reads at these sites to train a null model for hypothesis testing of the other sites. This is the intuition behind our rank-based filter procedure, which will now be described. Lastly we assume that read counts from distinct sites as well as distinct samples are statistically independent.

As stated above, the read counts $X_{i,j}$ for each position (i) and each replicate (j) are assumed to be independent for all i, j. By assumption, for each j, a subset of $X_{i,j}$ have the same distribution as NE mutants and therefore are identically distributed. We don't know which subset so our first goal will be to find an approximate subset that will have a distribution approximating the null distribution.

First we compute the rank of the read counts at each site within each replicate, averaging ties. Call the ranks $r_{i,j}$. Then, for each replicate, compute the mean rank of the *other* replicates (ie leave

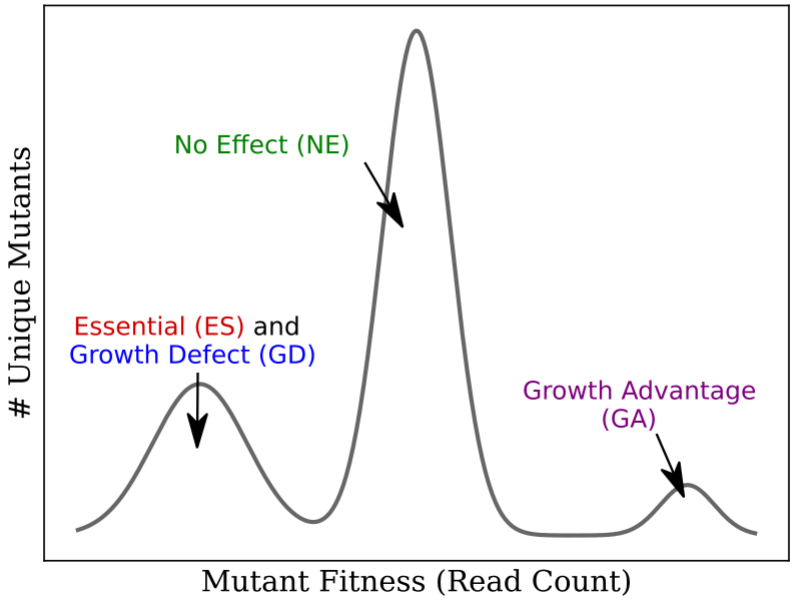

Figure 2.5: Conceptual diagram of the assumptions underlying our proposed rank-based filter method. We expect that most transposon mutants will not have a growth rate different from wild-type (ie are NE mutants). A substantial fraction will have a growth defect and a very small fraction will have a growth advantage. To estimate a null distribution a non-symmetric filter is necessary.

one out):

$$m_{i,j} = \frac{1}{J-1} \sum_{\tilde{j} \neq j} r_{i,\tilde{j}} \tag{2.1}$$

Then identify a subset of sites such that the mean ranks are within the expected 40% to 85% range:

$$S_j = \{i : 0.4 < \frac{m_{i,j}}{I} < 0.85\} \tag{2.2}$$

Finally, assemble the read counts of the null-distributed sites into the set $\bar{X}_j$, which is, by definition, a sample from an approximately null-distributed set of mutants.

$$\bar{X}_j = \{X_{i,j} : i \in S_j\} \tag{2.3}$$

Thus we have applied a rank-based filter to leave only a set of samples that are mutually independent and (approximately) null-distributed.

For simplicity, we index each element of the set $\bar{X}_j$ such that each read count is represented with the variable $Y_{k,j}$ for $k = 1, 2, 3, ..., K$. K is the number of insertion sites after applying the rank-based filter. Therefore $\{Y_{k,j} : k \leq K\} = \bar{X}_j$. $Y_{k,j}$ can now be used for fitting a zero-inflated negative binomial distribution (for ES identification) or for computing the ecdf (for GD/GA identification). Additionally, previous literature suggests that the Himar1 transposon is biased against insertion sites with the motif (GC)GNTANC(GC)[24]. Therefore, we separately apply the above rank-based filter to the read count data collected from these sites.

Hypothesis Testing for Essentiality. For simplicity, we will drop the j index as we perform identical calculations for each replicate using the corresponding $\bar{X}_j$. The zero-inflated negative-binomial distribution is:

$$N(y; \Theta, r, p) = \Theta \mathbf{1}[y = 0] + (1 - \Theta) \frac{\Gamma(r + y)}{y! \Gamma(r)} p^r \, (1 - p)^y \tag{2.4}$$

where $\mathbf{1}[...]$ is the indicator function, $\Gamma()$ is the gamma function. We will use maximum likelihood

estimation to fit the parameters. The log-likelihood is:

$$l(\Theta, r, p) = z\ln(\Theta + (1-\Theta)p^r) + (K-z)\ln(1-\Theta) + \sum_i \ln\left(\frac{\Gamma(y_i+r)}{y_i!\Gamma(r)}\right) + (K-z)r\ln(p) + \ln(1-p)\sum_i y_i \tag{2.5}$$

where K is number of samples and z is the number of samples that are zero (ie $z = \#\{y_k = 0\}$). The gradient is:

$$\frac{\partial l}{\partial \Theta} = z\frac{1-p^r}{\Theta + (1-\Theta)p^r} - (K-z)\frac{1}{1-\Theta} \tag{2.6}$$

$$\frac{\partial l}{\partial r} = z\frac{\ln(p)p^r(1-\Theta)}{\Theta + (1-\Theta)p^r} + \sum_i (\psi(y_i+r) - \psi(r)) + (K-z)\ln(p) \tag{2.7}$$

$$\frac{\partial l}{\partial p} = z\frac{(1-\Theta)p^{r-1}}{\Theta + (1-\Theta)p^r} + (K-z)\frac{r}{p} + \frac{1}{1-p}\sum_i y_i \tag{2.8}$$

where ψ is the digamma function. We solve for estimates of the parameters $(\hat{\Theta}, \hat{r}, \hat{p})$ with the L-BFGS-B algorithm (Scipy v1.2.1). We now precisely define an "essential" genes as one that, when removed, causes the bacterial colony to be 5% or less of the WT size (ie with a read count 5% of a mutant with no defect). While the particular threshold we have chosen is somewhat arbitrary, we feel it is both small enough to avoid misclassifying mutants as essential but not so small so as to have no hope of classifying highly defective mutants as ES. For hypothesis testing we define a "borderline ES" mutant by scaling our parameters such that the mean is 5% of the null model but the dispersion $(\frac{1}{r})$ is identical. Define $\tilde{p} = \frac{\hat{p}}{(1-0.05)\hat{p}+0.05}$. Thus, the cumulative distribution for a "borderline ES" mutant is:

$$F(y) = \sum_{\tilde{y}=0}^{y} N(\tilde{y}; \hat{\Theta}, \hat{r}, \tilde{p}) \tag{2.9}$$

Define a second function:

$$F^L(y) = \sum_{\tilde{y}=0}^{y-1} N(\tilde{y}; \hat{\Theta}, \hat{r}, \tilde{p}) \tag{2.10}$$

We calculate a continuously distributed p-value by sampling from a uniform distribution between F^L and F:

$$q_{i,j}^{ES} \sim U(F^L(X_i), F(X_i)) \tag{2.11}$$

where $q_{i,j}^{ES}$ is the p-value for site i, replicate j. We have included the second index (j) to emphasize that we will have a p-value for each replicate.

To pool essential p-values across samples we use the one-tailed Stouffer's method at each site:

$$q_i^{ES} = 1 - \Phi\left(\frac{1}{\sqrt{J}}\sum_{j=0}^{J}\Phi^{-1}(q_{i,j})\right) \tag{2.12}$$

Note that because low permissibility sites have a rather different distribution (far more likely to be zero) we fit a separate negative binomial for these special sites. We also fit separate distributions for each contig as sequencing depth varies greatly between contigs. P-values are fit in the same way as described above.

To pool p-values across insertion sites within a gene we use the truncated product method (TPM)[114] with a truncation threshold of 0.5 ($\tau < 0.5$). TPM provides a principled approach for limiting the effect of sites with no associated growth defect which would otherwise greatly inflate the p-values (such as those sites at the C-terminus of the gene which may not disrupt the function of the protein). We compute a p-value for gene g as:

$$\bar{q}_g^{ES} = \sum_{k=1}^{N_g}\binom{N_g}{k}(1-\tau)^{N_g-k}\left(w\sum_{s=0}^{k-1}\frac{(k\ln\tau - \ln w)^s}{s!}\mathbf{I}[w \le \tau^k] + \tau^k\mathbf{1}[w > \tau^k]\right) \tag{2.13}$$

where G_g is the set of sites belonging to gene g, N_g is the cardinality of G_g, and w is:

$$w = \prod_{i}^{N_g} q_i^{\mathbf{1}[q_i \le \tau]} \tag{2.14}$$

We then control the False Discovery Rate (FDR) using the Benjamini-Hochberg procedure (FDR $<$ 0.01)[9]. If the gene is rejected by this test it is declared ES.

GD/GA Hypothesis Testing. We utilize the read counts for insertion sites identified by the rank-based filter to form an approximate null distribution and use the ecdf to compute p-values. We define the ecdf for replicate j as:

$$H_j(y) = \frac{1}{K}\sum_{k=1}^{K}\mathbf{1}[Y_{k,j} \le y] \tag{2.15}$$

Also define:

$$H_j^L(y) = \frac{1}{K}\sum_{k=1}^{K} \mathbf{1}[Y_{k,j} < y] \tag{2.16}$$

where $H_j^L(0) = 0$. Note that because $Y_{j,k}$ takes discrete values, $H_j^L(y)$ and $H_j(y)$ will generally differ. To calculate a p-value for site i (replicate j) we sample a uniform distribution bounded by these two values:

$$q_{i,j}^{GD/GA} \sim U(H^L(X_i), H(X_i)) \tag{2.17}$$

As before we will compute a separate ecdf for low permissibility sites and for each contig in the genome.

For a particular insertion site, the p-values from each sample are pooled using the one-tailed Stouffer's method, as before. The resulting pooled p-values from all insertion sites within the same gene are then pooled using a two-tailed version of Stouffer's method.

$$\bar{q}_g^{GD/GA} = 2\left(1 - \Phi\left(\frac{1}{\sqrt{N_g}}\left|\sum_{i\in G_g} \Phi^{-1}(q_i^{GD/GA})\right|\right)\right) \tag{2.18}$$

Family-wise error rate (FWER < 1.0) using the Bonferroni correction (ie only a single falsely detected GA/GD is expected in the results).

Relative fitness (RF). The mutant fitness, relative to wildtype, resulting from disruption of a particular gene is approximated as follows. First, the mean of the read counts at each insertion site is calculated across samples:

$$\mu_i = \frac{1}{J}\sum_{j}^{J} X_{i,j} \tag{2.19}$$

The site fitness is calculated as the mean read count of each site divided by the median across sites (ie samples are normalized to the median, which is assumed to have no growth defect):

$$f_i = \mu_i / \underset{i}{\text{med}}\, \mu_i \tag{2.20}$$

Finally, each gene is assigned a Relative Fitness equal to the median of the site fitness for all sites contained in the gene g.

$$RF_g = \underset{i\in G_g}{\text{med}}\, f_i \tag{2.21}$$

Table 2.2: Table of features annotated by our analysis method. NE = No Effect, GD = Growth Defect, ES = Essential, GA = Growth Advantage, N/A = Feature lacks potential insertion sites (TA dinucleotide) for the Himar1 transposon or only contains sites shared with another feature.

	CDS	**Pseudogene**	**tRNA**	**Riboswitch**	**rRNA**	**ncRNA**	**tmRNA**
NE	2850	117	8	5	2	1	0
ES	259	0	8	0	2	0	1
GD	1208	32	26	0	0	1	0
GA	460	29	0	0	0	0	0
N/A	64	13	4	1	0	0	0

A gene is declared GD if its Relative Fitness is less than 2/3 and is statistically significant after controlling FWER. Similarly, a gene is declared GA if its Relative Fitness is greater than 1.5 and is statistically significant after controlling FWER.

2.4.3 Results

Detection of essential MAC109 genetic features. Our analysis method classified 270 features as ES, 489 features as GA, 1267 features as GD out of 5091 total annotated features. 73 features contained zero thymine-adenine dinucleotides (TA sites) and 9 features only contained TA sites shared with another feature. These 82 features therefore cannot be evaluated with the Himar1 system which only transposes into TA sites. A summary of classifications by feature type is provided in Table 2.2. Genes classified as essential by our method are provided in Appendix A. Interestingly, our method identified 3 annotated coding sequences in pMAC109a and 1 coding sequence in pMAC109b as essential.

Comparison of essential genes with previously published results. We also compared the results of our analysis method applied to a previously published Tn-seq dataset using Mtb strain H37Rv[24]. All genes labelled as "ESD" (containing an essential domain) in the previously published dataset were considered essential for comparison. Figure 2.6 shows the overlap in the predicted essential coding sequences (CDS) from each method (RNA and other features excluded). Overall there was good agreement between each method though our method appears to be somewhat more sensitive for essential gene detection than the HMM-based method at this sample size. Upon inspection it was observed that the essential genes unique to our method contained a significant number of sites with zero or very few insertions, but these low read sites were interspersed with sites

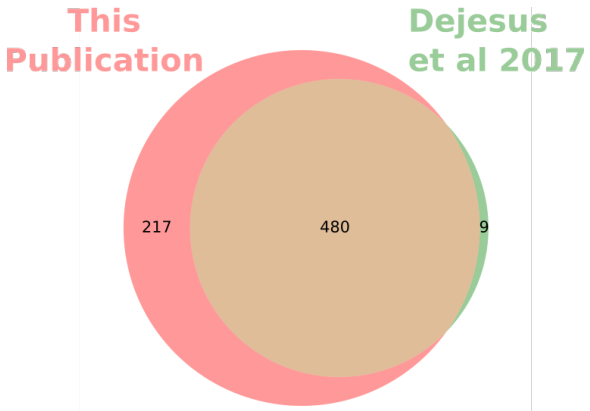

Figure 2.6: Venn diagram of essential genes predictions for Mtb strain H37Rv from our analysis compared to the previously published essential gene predictions from Dejesus et al [24]. Notably, the genes labelled essential by the HMM are nearly a subset of the genes labelled as essential by our method. Only protein coding sequences are considered in this diagram.

containing larger numbers of reads. This fits with expectations that the HMM model is sensitive primarily to multiple adjacent sites with low read counts, whereas our method is sensitive to the number of sites per gene regardless of adjacency. The full list of overlapping essential genes is given in appendix E.

2.5 Summary and Discussion

This chapter discussed our efforts to characterize a frequently utilized patient-derived strain of Mav. This is intended to provide some of the initial information necessary for developing interventions to prevent, diagnose, and treat Mav infections. We provided a high-quality and complete genome along with annotation information. This included gene finding and Ontologies which we expect will be most helpful for the interpretation of future high-throughput sequencing data. Additionally, we optimized the transposon mutagenesis of this organism and Tn-seq library preparation enabling us to collect a high-quality Tn-seq dataset. We developed a novel bioinformatics method for identifying essential genes from this data which utilizes fewer assumptions about the distribution of read counts - leading to potentially more reliable results. Finally, we provided a list of the essential genes of Mav which we expect will be highly useful for prioritization of drug targets and for further characterization of Mav genes. Incidentally, we found apparently essential genes in the plasmids of Mav. We discuss

some possible explanations for this surprising result in chapter 5.

Chapter 3

Adjunctive drug targets in *Mycobacterium avium*

3.1 Objectives

As previously discussed, treatment of infections with Mav requires years of multi-drug therapy. The ability of the bacterium to survive in the presence of antibiotics for such a lengthy period (ie persist) likely requires the action of bacterial genes. There are two non-exclusive possibilities for the organization of the bacterial persistence response. First, it is possible that upon treatment with many different antibiotics that a single bacterial "persistence pathway" is induced which allows the bacteria to survive multiple drugs. Secondly, it is possible that there are multiple, antibiotic-specific persistence mechanisms. To investigate how Mav is able to persist in the presence of antibiotics we utilized our constructed genome-wide transposon mutant pool (see Ch. 2) to do a high-throughput screen. Our hypothesis was that there are non-essential genes (or other other genetic elements) involved in the persistence phenomenon. We reasoned that a transposon mutant lacking a gene in the persistence pathway should be hypersusceptible to antibiotics which could be detected using the Tn-seq library prep protocol. Characterizing the role of such a gene might suggest new avenues for targeting persisters and enable improved, shorter therapies for treating infections with Mav.

Additionally, the constructed mutant pool also provided an opportunity to explore the mechanisms of antibiotic action. While the initial steps by which major antibiotics kill bacteria have largely been elucidated, important unanswered questions remain. Generally speaking a particular antibiotic will bind to and inhibit the function of a certain target enzyme. For example the drug rifabutin will

reversibly bind to and inhibit bacterial RNA polymerase. However, there remains significant uncertainty about the sequence of events leading from the antibiotic binding event to bacterial death. We hypothesized that there were non-essential genes involved in the "death pathway" triggered by antibiotic binding in mycobacteria. Transposon mutants with defects in this hypothetical pathway would appear in a screen to be *hypertolerant* to antibiotics (ie would survive better than wildtype). Identifying these mutants and characterizing the function of the mutated gene could lead to a deeper understanding of the mechanisms of antibiotics and suggest avenues for therapy.

3.2 Differential susceptibility screen

3.2.1 Drugs

CLR, RFB, moxifloxacin (MOX), and EMB were used in these screens. These drugs were chosen as CLR, RFB, and EMB are commonly used in the treatment of Mav. MOX may be useful in treating patients who fail macrolide-containing regimens [58]. Ranges of drug concentrations were chosen based on estimates of the maximum accumulated doses in human lung tissue (likely the highest dose the bacteria will experience). Dilutions from this maximum dose were also included to simulate reduced levels of drug, which may be more realistic. Doses achieved in lung tissues were taken to be 54ug/mL for CLR[31], 0.63ug/mL for RFB[10], 10.0ug/mL for MOX[83], and 21.0ug/mL for EMB[66] (based on non-human primate data). Final ranges of drug concentrations were a 10-fold dilution series for each compound: CLR (0.54 - 54 ug/mL), RFB (0.0063 - 0.63 ug/mL), MOX (0.1 - 10.0 ug/mL), EMB (0.21 - 21.0 ug/mL). All drugs were dissolved in dimethyl sulfoxide (DMSO). 3 biological replicates were performed for each drug at each concentration. As described below, only samples using two doses of each drug were processed for sequencing.

3.2.2 Conditions for screen

The setup for our genome-wide differential susceptibility screen is sketched in figure 3.1. We inoculated a 1mL aliquot of the combined transposon mutant pool (consisting of approximately 1.2×10^6 unique mutants) into 300mL 7H9/30%OADC contained in a 1.3L roller bottle. This was shaken at 37°C for 24 hours to reduce bacterial clumping (220rpm, 0.75in orbit). The optical density (OD)

was tracked until exceeding 0.32. The transposon mutants were then diluted to OD 0.1 with cold 7H9/30%OADC and aliquoted to 89 50mL conical tubes, 10mL/tube. Tubes were incubated for 5 hours shaking at 37°C to allow the bacteria to return to log-phase. 5 tubes were randomly selected for processing (0 hours). These samples were then processed for CFU enumerate and regrowth (see below for additional processing details). Note that regrowth is required in order to remove the contribution of DNA from dead transposon mutant bacteria. Drugs (or DMSO) were then added to all tubes. Samples were taken at 12 and 48 hours after drug addition and processed for CFUs and regrowth in the same manner.

Sample processing for CFU enumeration and regrowth. CFUs were estimated throughout by removing 400uL of bacterial culture, centrifuging (2000g for 5 minutes) and washing twice with PBS-Tw (to remove antibiotic). 10-fold dilutions were done in PBS-Tw and 50uL of each dilution was plated on 7H11 agar. T-shaped spreaders were used to spread liquid evenly across plates. CFUs were counted after 7-8 days. For regrowth, the remainder of each tube was twice centrifuged (2000g for 10 min) and washed with 10mL PBS-Tw to remove drug. Centrifugation was done a final time and the sample was reimmersed in 250uL PBS-Tw. 50uL of the washed transposon pool was plated on each of four 7H11 agar plates and spread with 10-15 3mm sterile glass beads. Samples were regrown for 7-8 days. Bacterial lawns for the four agar plates were scraped and pooled into 2mL tubes. DNA extraction was done on regrown samples as described in 2.2 for short read sequencing. DNA was processed for Tn-Seq as described in appendix B. Libraries were sequenced (2×75bp) on an Illumina HiSeq 2500 by the Johns Hopkins GRCF High Throughput Sequencing Center. A total of 59 samples (5 input pool plus 18 groups of triplicates) were sequenced yielding between 2,333,295 – 7,193,522 reads per sample for a total of 269,324,560 paired-end reads.

3.2.3 Mutant hypersusceptibility validation

Himar1 transposon mutants in the orthologs of DFS55_00120 (B6K05_00330), DFS55_00360 (B6K05_00550), and DFS55_12665 (B6K05_12310) in the MAH11 [113] background were the kind gift of Marte Dragset (NTNU). These were used to validate our predicted antibiotic differential susceptibility predictions. Before testing we confirmed the location of the transposon mutant using a previously

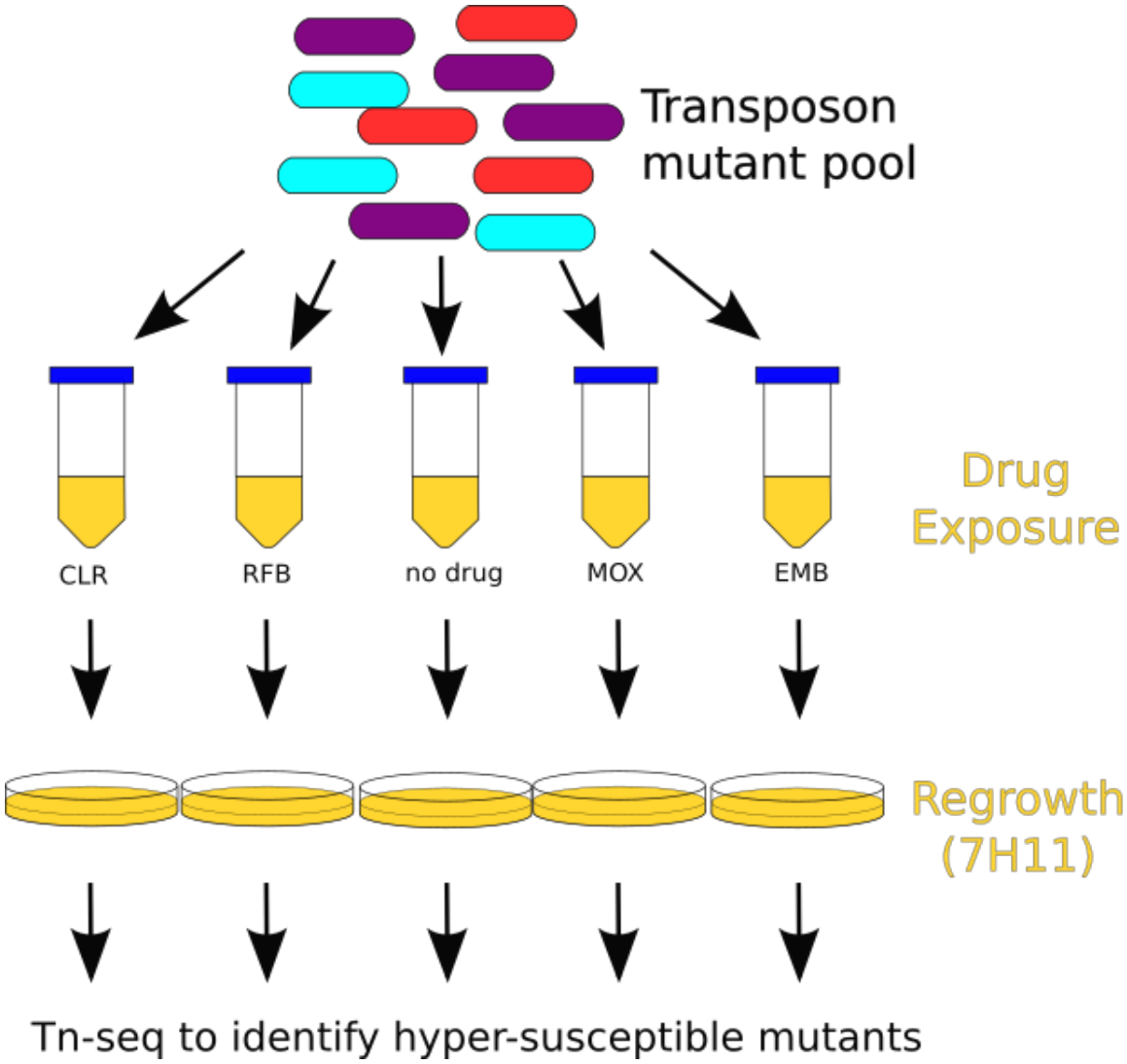

Figure 3.1: Schematic of experimental setup for identifying adjunctive drugs. A transposon mutant pool constructed in MAC109 (see Ch.2) is grown in a single vessel and aliquoted to tubes. After a period of adjustment to the new conditions antibiotics (CLR,MOX,RFB,EMB) or vehicle control (DMSO) are added. After 12 and 48 hours of drug treatment samples are taken for regrowth and CFU enumeration.

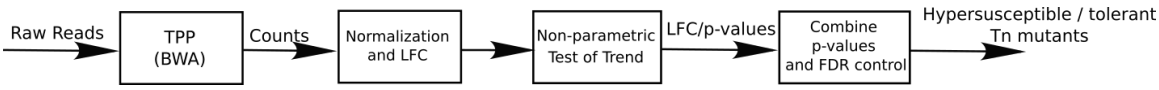

Figure 3.2: Schematic of analysis pipeline for processing Tn-seq data.

published protocol [81]. Locations of insertions were at bases 186/468 (186bp from start of gene, total gene length 468 bp), 1504/2634, and 546/2334 for DFS55_00120, DFS55_00360, and DFS55_12665, respectively.

Mutants were tested for hypersusceptibility by growing each strain (and MAH11 wildtype) in 7H9/30%OADC. After strains reached an OD of more than 1.0 they were filtered through a 5um syringe filter to remove clumps, as previously described [19]. We found that using a somewhat high OD greatly reduced the clumping of *Mav* which permitted more accurate enumeration of colony forming unitss (CFUs) (via light microscopy, data not shown). Cultures were then diluted to OD $\sim$ 0.07 in 10mL of fresh 7H9/30%OADC in 50mL conical tubes and allowed to regrow to OD $>$ 1.0. Drugs were then dissolved in DMSO and added. Final drug concentrations for validation were 54ug/mL CLR, 10 ug/mL MOX, 21 ug/mL EMB. DMSO was added to no-drug control. 3 replicates were performed for each strain-drug pair chosen for validation (33 tubes total). Samples were taken 2 and 4 days after adding antibiotics and diluted in 7H9/30%OADC for CFU enumeration on 7H11 agar plates.

3.3 Computational analysis

A schematic of the pipeline to process the data is provided in figure 3.2. Briefly, raw reads were mapped as in 2.4.1 using tpp. Counts from tpp were then processed with a custom python script (Python 3.7) to produce a *.csv file to be read by pandas 0.24.1.

Effect size/log fold change calculation. For calculation of log fold change (LFC) a pseudocount of 4 reads was added to all samples before dividing read counts by total read count:

$$\tilde{x}_{t,ir} = \frac{X_{t,ir} + \alpha}{\sum_t^T (X_{t,ir} + \alpha)} \tag{3.1}$$

Where $X_{t,ir}$ is the read count for insertion site t, for antibiotic treatment group i, for replicate r. α is the pseudocount ($\alpha = 4$) and T is the number of insertion sites. A large pseudo count used to

stabilize the fold-change which can vary dramatically due to mutants with low representation. The stabilized-normalized read counts were then averaged:

$$\mu_{t,i} = \frac{1}{n_i} \sum_{r}^{n_i} \tilde{x}_{t,ir} \tag{3.2}$$

where n_i is the number of treatment groups, and LFC between sample group i and j calculated as the median of the LFC at individual sites with a gene:

$$\mathrm{LFC}_{g,(i/j)} = \mathrm{med}_{t \in G_g} \left(\log_2 \left(\frac{\mu_{t,i}}{\mu_{t,j}} \right) \right) \tag{3.3}$$

where G_g are the set of sites belonging to gene g and $LFC_{g,i/j}$ is the log-fold change between experimental groups i and j for gene g.

P-value calculation. For calculation of p-values, read counts for each sample were first normalized by dividing by the total read count in each sample (no pseudocounts were used).

$$x_{t,ir} = \frac{X_{t,ir}}{\sum_t^T X_{t,ir}} \tag{3.4}$$

The Jonckheere-Terpstra (JT) test was applied at each time point[55, 97]. Briefly, the JT test is a non-parametric test of trend which is more powerful than the more common Kruskal-Wallis test when the alternative hypothesis assumes a (monotonic) trend of the groups. In this case we have three groups for each drug at each time point: No drug, low dose, high dose. We expect that if a mutant is hypersusceptible at a low dose it will be even more hypersusceptible at a higher dose. We define the alternative hypothesis as:

$$H_A : \theta_1 \leq \theta_2 \leq ... \leq \theta_K \tag{3.5}$$

against a null hypothesis:

$$H_0 : \theta_1 = \theta_2 = ... = \theta_K \tag{3.6}$$

where $\theta_1 ... \theta_K$ are measures of a centrality parameter (such as the median) and K is the number of treatment groups for the drug at a particular time point (in this case $K = 3$ for all groups). The

JT test statistic at site t is defined as:

$$B_t = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \sum_{r=1}^{n_i} \sum_{s=1}^{n_j} \mathbf{1}[x_{t,ir} < x_{t,js}] \tag{3.7}$$

Computation of p-values at individual sites are computed via permutation test. Naively this would result in non-uniformly distributed p-values due to the discrete nature of the distribution. To avoid this we perform uniform random sampling between the two bounding cumulative distribution functions (cdfs) as described in equation 2.17. Pooling the p-values within each gene is then accomplished with the two-sided Stouffer's method (see equation 2.18). Finally, adjusted p-values are then computed using the Benjamini-Hochberg procedure. Gene mutants are considered "differentially susceptible" to a drug if the absolute value of the $\log_2$-fold change (relative to no drug control) is greater than 0.5 and the adjusted p-value is less than 0.05 at both 12 hours and 48 hours.

3.4 Results

CFUs for pools at 0h, 12h, and 48h are provided in figures 3.3 - 3.6. The same no-drug (DMSO-only) control appears in all plots. Notably the no-drug control has an inflection at the 12h timepoint. This is likely due to the clumping of the bacteria soon after starting cultures. Under a light microscope (unstained) the no-drug control cultures showed clumps of approximately 5 cells. It is likely clumping was also present in the drug containing tubes, though these were not examined. Despite the clumping, which prevents a clear interpretation of the CFU curves, the screen should not be impacted as each bacterium within each clump will have an identical chance of regrowth. Thus our major results (the differential representation of mutants exposed to drugs) should not be impacted by the presence of clumping in our cultures.

Only a subset of samples were chosen to proceed with Tn-Seq. We considered the following criteria of an antibiotic exposure: (1) Bacterial numbers must exceed 10^6 CFU/mL at all times during exposure - this ensures that the probability of losing a particular mutant is minimized. (2) There should be detectable attenuation of bacterial viability after adding the drug - this ensures the concentration of antibiotic is high enough to remove highly-susceptible mutants. Otherwise the concentration might be too low to select against these mutants. (3) A preference for drug

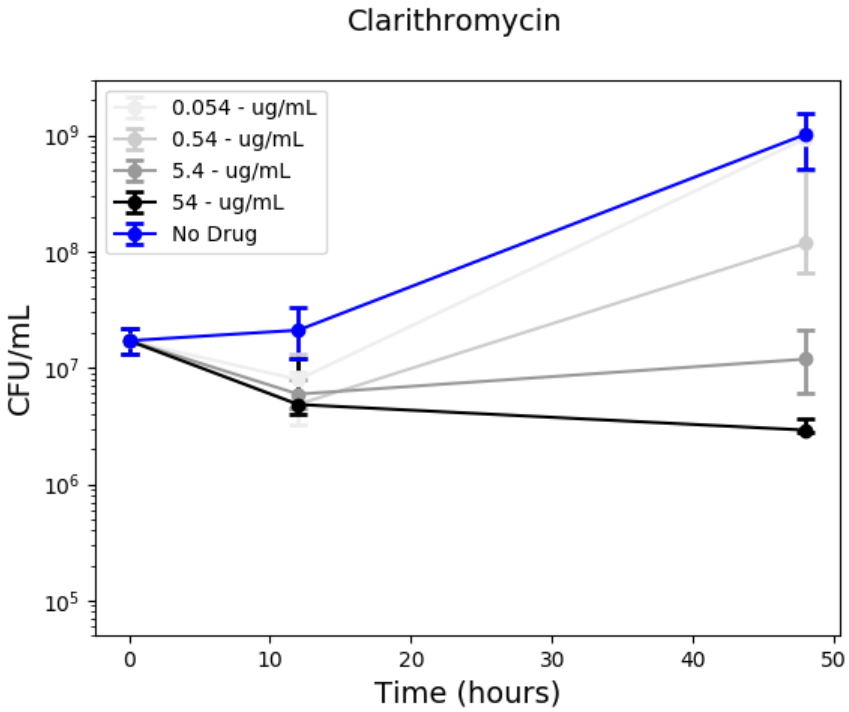

Figure 3.3: Bacterial viability of transposon mutant pools after addition of CLR at 0 hours. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

concentrations near or below achievable serum concentration of drug after standard doses. We take serum values as 2.31 ug/mL CLR, 4.42 ug/mL MOX, 0.52 ug/mL RFB and 2.27 ug/mL EMB [101]. Using this set of criteria we decided to sequence both time points at the following concentrations: CLR 0.54 and 5.4 ug/mL, MOX 0.1 and 1.0 ug/mL, RFB 0.063 and 0.63 ug/mL, EMB 2.1 and 0.21 ug/mL.

3.4.1 Hyper-susceptible mutants

We identified 67 mutants as differentially susceptible for CLR (50 hypersusceptible), 9 for EMB (9 hypersusceptible), 109 for MOX (106 hypersusceptible), 104 for RFB (82 hypersusceptible). Effect sizes for these mutants are plotted in figures 3.7 - 3.10. Significant mutants and effect sizes are also provided in appendix C. Additionally, 3 mutants were found to be differentially susceptible (all were hypersusceptible) to all four drugs including DFS55_00905 (annotated as “acyltransferase”), DFS55_10120 (MoxR family ATPase), and DFS55_12730 (hypothetical protein).

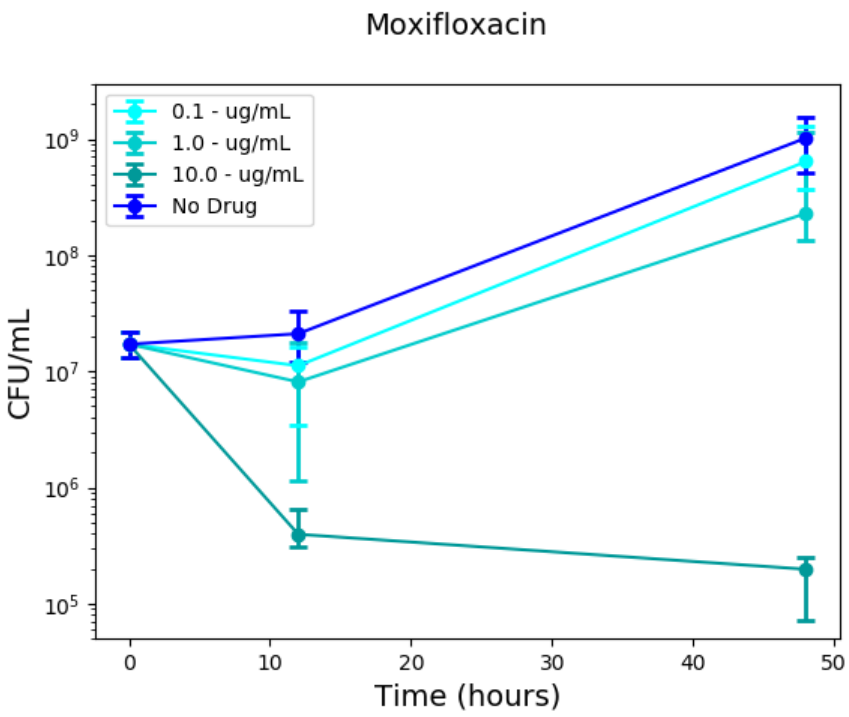

Figure 3.4: Bacterial viability of transposon mutant pools after addition of MOX at 0 hours. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

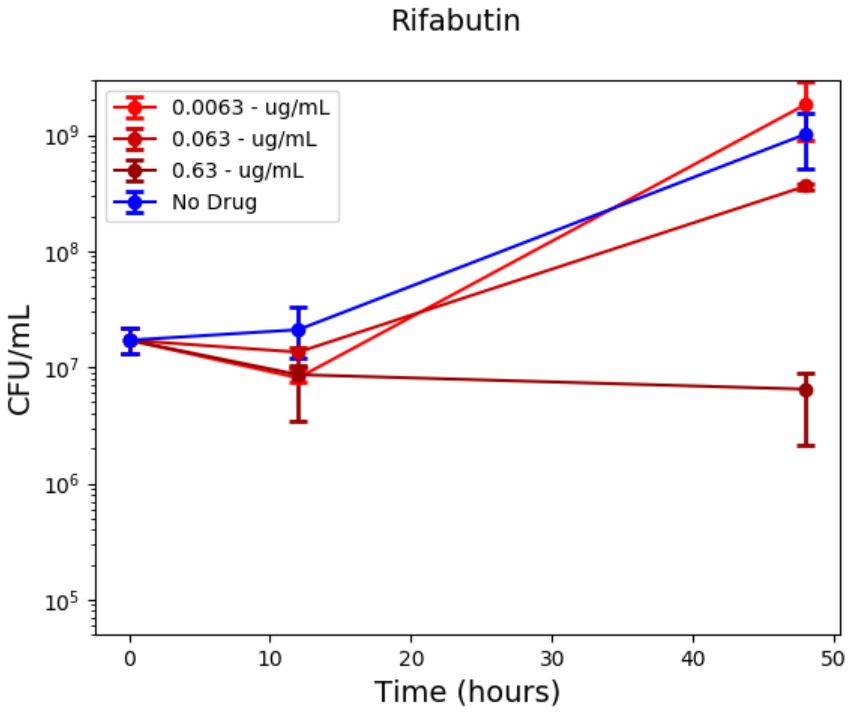

Figure 3.5: Bacterial viability of transposon mutant pools after addition of RFB at 0 hours. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

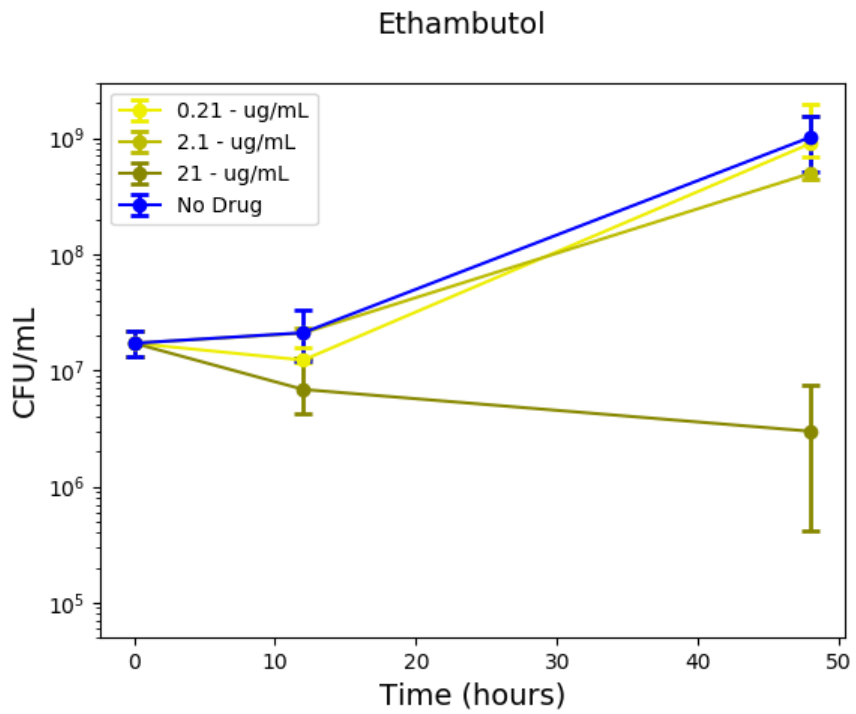

Figure 3.6: Bacterial viability of transposon mutant pools after addition of EMB at 0 hours. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

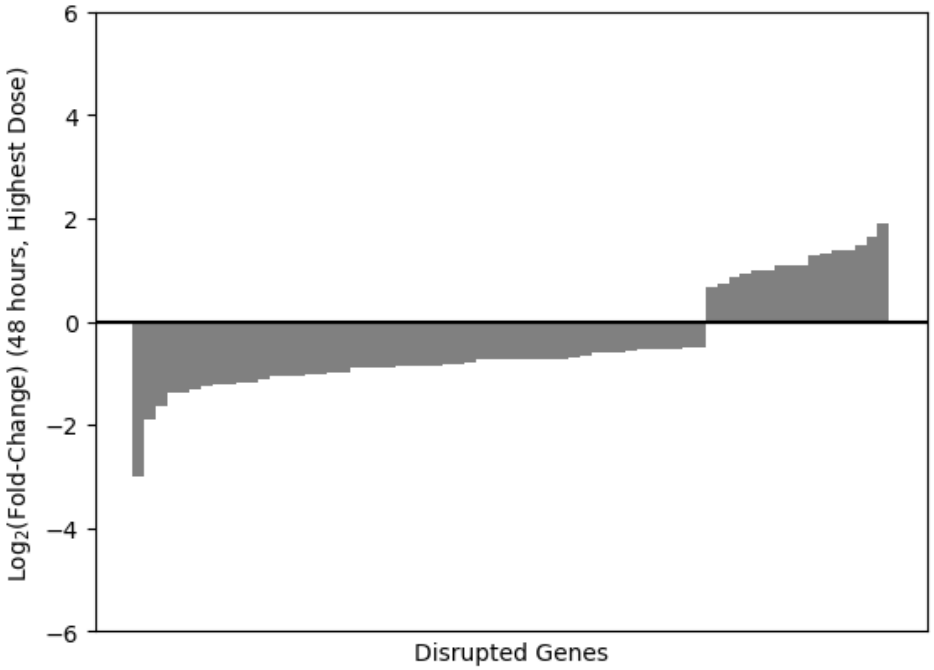

Figure 3.7: Effect size of disruption of genes in MAC109 (fold-change is normalized to no drug control). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to CLR, positive values are hyper-tolerant ones.

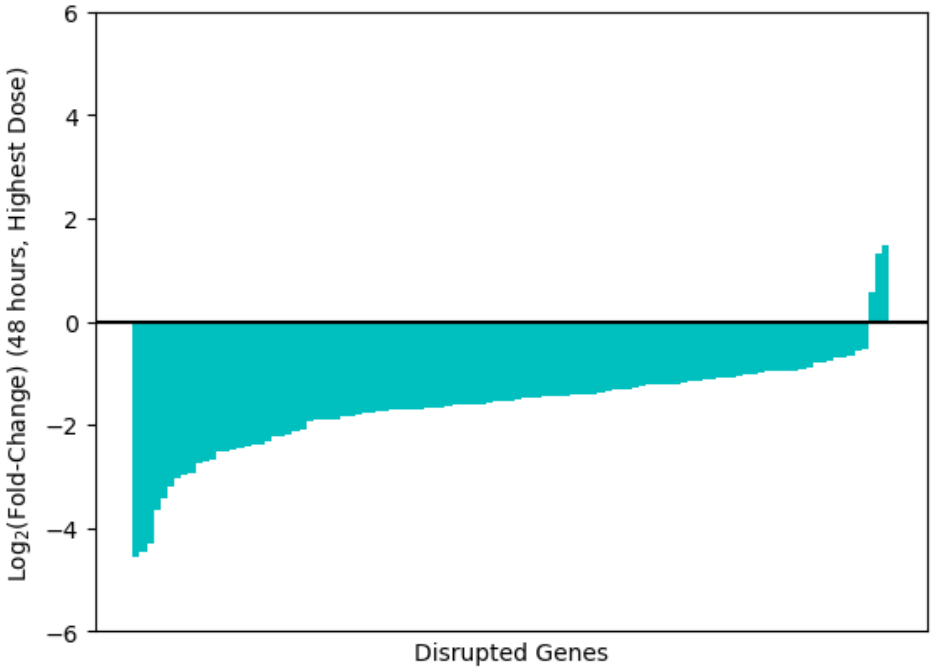

Figure 3.8: Effect size of disruption of genes in MAC109 (fold-change is normalized to no drug control). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to MOX, positive values are hyper-tolerant ones.

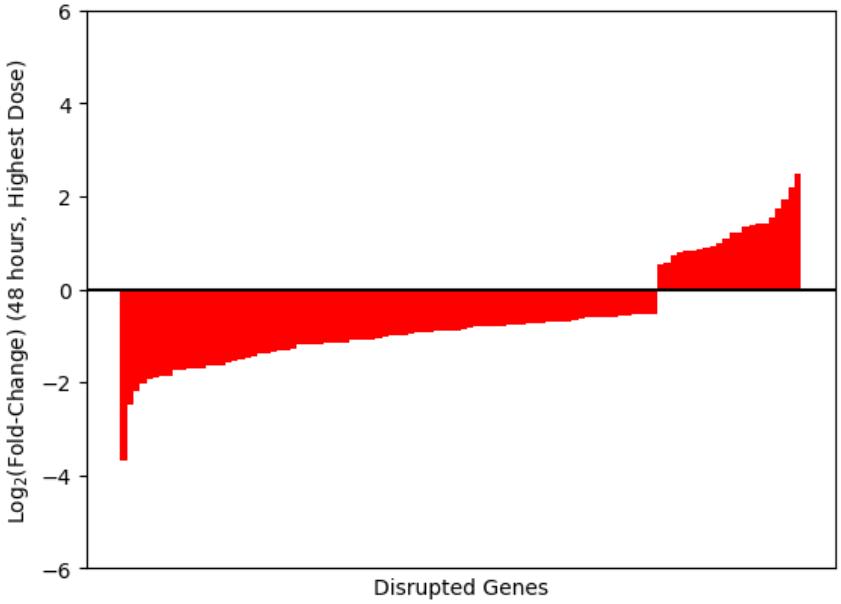

Figure 3.9: Effect size of disruption of genes in MAC109 (fold-change is normalized to no drug control). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to RFB, positive values are hyper-tolerant ones.

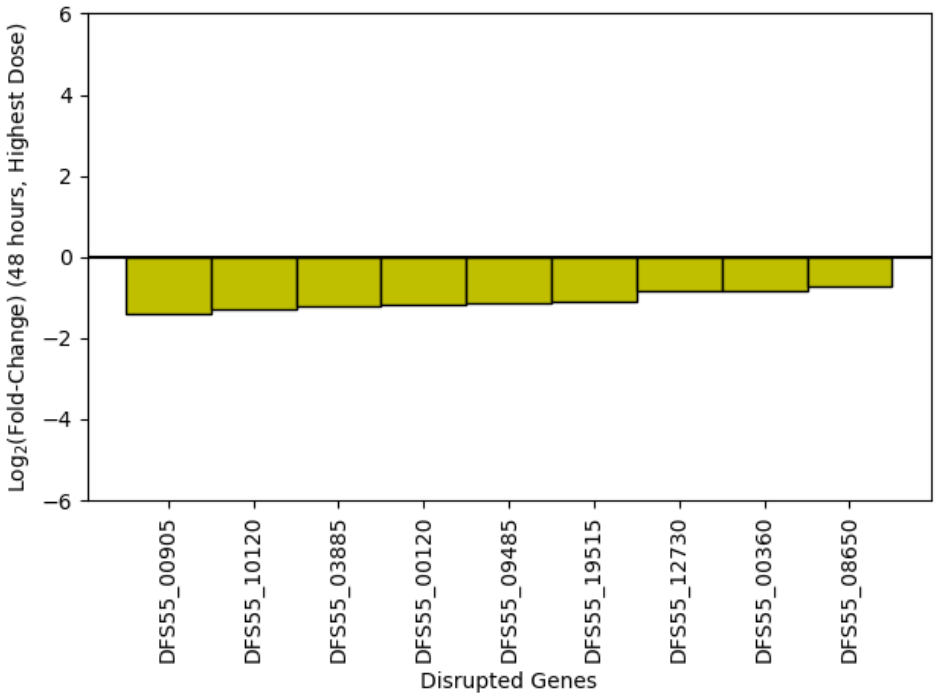

Figure 3.10: Effect size of disruption of genes in MAC109 (fold-change is normalized to no drug control). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hypersusceptible to EMB, positive values are hyper-tolerant ones. Names of genes (DFS55_#####) given at bottom.

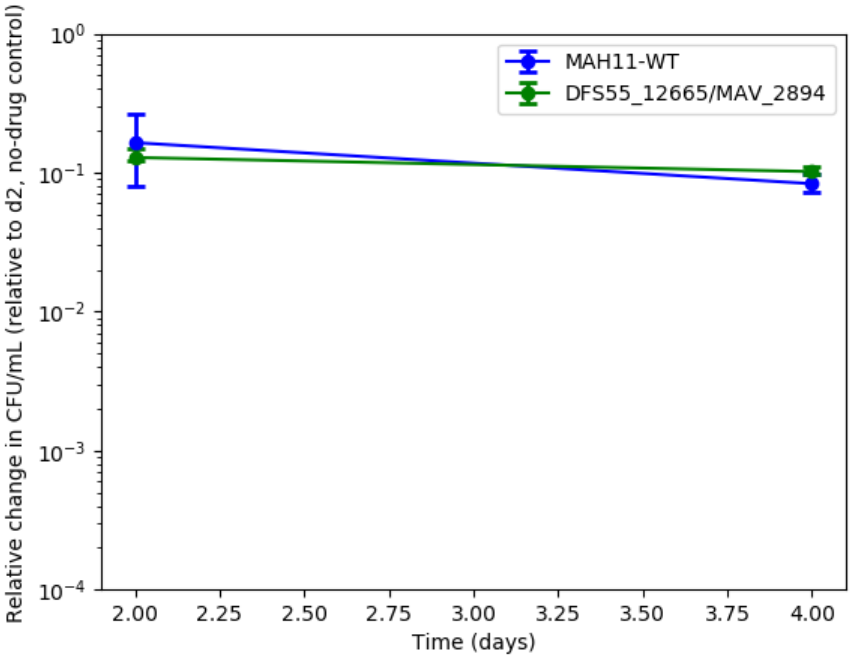

Figure 3.11: Plotted is the change in viability of each strain to CLR normalized to the no-drug control (DMSO) at 2 days (48h) and 4 days (96h). Smaller values indicate greater susceptibility to the antibiotic.

3.5 Validation of hypersusceptible mutant phenotypes

To validate our predicted antibiotic hypersusceptibilities we utilized isolated transposon mutants in the genes DFS55_00120 (predicted to be hypersusceptible to EMB), DFS55_00360 (hypersusceptible to EMB), and DFS55_12665 (hypersusceptible to MOX and CLR). The annotations were "FHA domain-containing protein", "penicillin-binding protein", and "accessory Sec system translocase SecA2", respectively. To compare results between strains we normalized to the median of day 2 CFUs of no drug controls (for each strain). Our results confirmed hypersusceptibility for EMB and MOX but we did not observe hypersusceptibility to CLR. Notably, hypersusceptiblity of DFS55_00120 to EMB was much greater than DFS55_00360.

3.6 Summary and Discussion

In this chapter, we utilized a genome-wide transposon mutant pool in Mav (created as described in chapter 2) to screen for mutants with a growth phenotype in the presence of antibiotics. Our

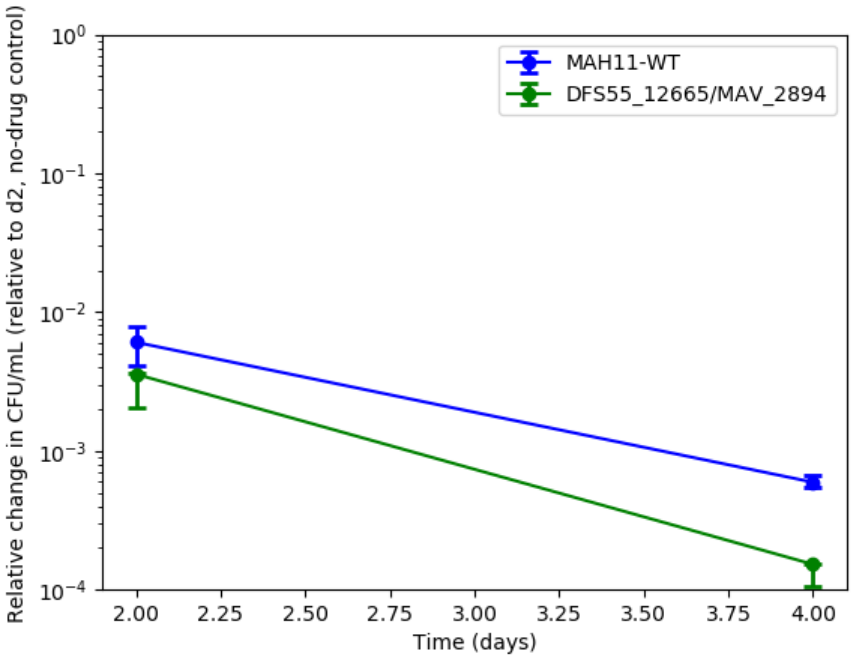

Figure 3.12: Plotted is the change in viability of each strain to MOX normalized to the no-drug control (DMSO) at 2 days (48h) and 4 days (96h). Smaller values indicate greater susceptibility to the antibiotic.

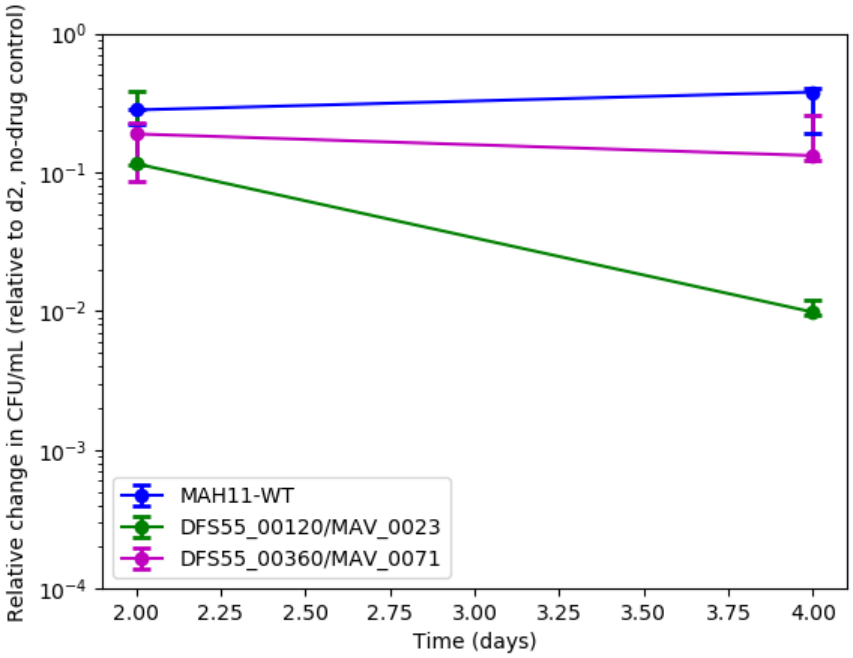

Figure 3.13: Plotted is the change in viability of each strain to EMB normalized to the no-drug control (DMSO) at 2 days (48h) and 4 days (96h). Smaller values indicate greater susceptibility to the antibiotic.

hypothesis was that mutants with growth phenotypes might be helpful for a) understanding the mechanisms of bacterial persistence in Mav and b) identifying the events that occur leading up to bacterial death in the presence of antibiotics. The hope is that this information might be useful for shortening regimens used to treat Mav.

There were limitations with our approach. First, mutants in essential genes could not be screened as these cannot be constructed using transposon mutagenesis. Therefore essential genes that play a role in these processes will be missed. Second, gene disruptions leading to changes in extracellular factors (eg. extracelluar proteins) may be missed by our screen as these factors may be complemented by factors produced by non-defective mutants in the same tube.

Using our screen and high-throughput sequencing (Tn-Seq) we identified hundreds of mutants likely to have a growth phenotype in the presence of antibiotics. MOX exposure yielded predominantly hypersusceptible mutants with only three hypertolerance mutants. This may suggest that there are many potential targets to boost the efficacy of moxifloxacin and that the pathway leading from inhibition of topoisomerases by moxifloxacin to bacterial death only involves a few genes. Additionally, given the strong effect sizes observed with moxifloxacin relative to the other drugs this antibiotic may represent the greatest opportunity for adjunctive therapy to boost its activity.

We followed-up a few of our hypersusceptibility results using individual isolated transposon mutants. It is important to note that there were differences between the conditions of our screen and the follow-up experiment. The individual mutants were isolated in MAH11 - an entirely separate isolate from the MAC109 strain used during our screen. This was necessary as a collection of mutants in the MAC109 strain is currently unavailable. The relatively small size of the available library (2-3 thousand mutants) also limited the mutants we were able to select for follow-up. Follow-up work should select mutants with the greatest measured effect size assuming they are available or can be easily generated. Secondly, we utilized a late-log to early stationary phase culture for validation whereas our original screen was performed with log-phase cultures. Using a late-log culture was found to greatly reduce bacterial clumping. This allowed us to more easily interpret the CFU counts - our measure of hypersusceptibility. However, this may have impacted the state of the bacteria,

modifying the killing effect of the antibiotics. Given the differences in setup it is reassuring that were able to confirm hypersusceptibility for EMB and MOX among the three mutants tested. However, the failure to validate the predicted hypersusceptibility to CLR of the DFS55_12665 mutants does raise questions about the reliability of our predictions derived from the Tn-seq data. Follow-up work should seek to quickly validate predictions, preferably through CFU counts as reported here, before further investigations are carried out.

Pending additional validation, some preliminary conclusions can be made. The hypersusceptibility of the DFS55_12665 mutant to MOX but not CLR suggests that hypersusceptibility is specific to a particular drug. This may mean that antibiotic persistence is a drug-specific phenomenon, meaning a persister to one drug may not necessarily be a persister to another. If validated this would suggest that highly persistent infections might best be treated using a combination of multiple classes of antibiotics in order to target the different types of persisters. Future work might focus on the study of persistence to a single particular drug as opposed to pan-drug persistence.

Chapter 4

Adjunctive drug targets and mechanisms of antibiotic tolerance in Mtb

4.1 Objectives

Chapter 3 described out efforts to identify genes involved in the response of Mav to antibiotics. Similarly, we wished to identify mechanisms by which the global pathogen Mtb is able to survive exposure to antibiotics as well as those genes involved in our hypothesized antibiotic "death pathway". Here we construct genome-wide transposon mutant pools in Mtb and perform similar screens to those performed in Mav to explore these questions. This also provides us the opportunity to compare the identified genes from Mtb with those identified in chapter 3 for Mav, which may help to identify general mycobacterial genes affecting antibiotic killing.

Additionally, we wanted to explore the role of the bacterial extracellular environment in the acquisition of the persistence phenotype. We hypothesized that the pathways involved in antibiotic survival are specific to the environment. In particular we were interested in environments known to greatly reduce antibiotic effectiveness (such as PBS[36, 112]). Identifying mutants with hypersusceptibility phenotypes may be particularly useful for understanding the mechanisms by which these environments induce an antibiotic persistent phenotype, which are likely connected with the massive change in metabolism that occurs during starvation [40, 36]. These discoveries may be most relevant to the environment that Mtb experiences during host infection, which have been speculated to have

limited nutrient availability [68]. Furthermore, if such genes can be identified and characterized they may make highly effective drug targets for shortening TB therapy.

4.2 Experimental Methodology

4.2.1 Strains, medium, and buffers

PBS-Tw, MP Buffer and 7H11 agar were made as described in chapter 2. To make 7H9/Tw+/G+/10% OADC: 2.35g 7H9 powder was added to 450mL deionized water. 2mL of 50%glycerol was then added. After sterilization (via autoclaving at 121°C or by passing through a 0.22um filter) and cooling to room temperature, 50mL of OADC enrichment (Becton-Dickinson) and 1.25mL 20% Tween-80 were added. To make 7H9/Tw+/G-/10% OADC: Same as above excluding glycerol.

Transposon mutant pools were constructed in H37Rv by adapting a previously published protocol [69] using ϕmycomarT7. Mtb strain H37Rv was grow from frozen stocks in 450mL 7H9/Tw+/G+/10% OADC to OD 1.2 in a large roller bottle at 37°C, shaking. Culture was split into 7 tubes of 50mL/tube. Tubes were centrifuged (2000 g for 5 min) and resuspended in 10mL MP Buffer. This washing step was repeat 2 additional times. Cells were then spun an additional time (2000g for 5 min) and reimmersed in 4mL MP Buffer. $\sim 1 \times 10^{11}$ PFUs (10:1 phage:bacilli) of ϕmycomarT7 was added to 6 of the tubes - the seventh tube recieved only MP Buffer without phage. Tubes were place on a shaking incubator (37°C) for two days. After incubation the transformation mixture was centrifuged (2000g for 5 min) and washed with PBS-Tw to remove residual phage. An additional centrifugation was done and bacteria were reimmersed in 1mL PBS-Tw.

50uL of each tube of washed transformants (or no-vector control) were diluted and plated on 7H11 plates, with or without 50ug/mL kanamycin, to determine transformation efficiency and background resistance. The remainder of the cultures were plated on 7H11 containing 50ug/mL kanamycin in Pyrex baking dishes (15" x 10", 500mL agar per dish, 1 tube per dish). After 35 days colonies were scraped from each dish and dispersed by vortexing with sterile glass beads (3mm) in fresh 7H9 broth and frozen in aliquots at -80°C for later use. For high-throughput screens equal volumes of all six pools were combined and frozen in aliquots to increase mutant diversity.

Strains for validation. The CDC1551, caeA-KO mutant, and caeA-COM strains were the

kind gift of Dr. Shichun Lun and the laboratory of Dr. Bill Bishai. As previously described [70], caeA-KO is a knockout strain of the caeA/Rv2224c in the CDC1551 background. caeA-COM is the caeA-KO strain complemented via an integrating plasmid with an intact copy of the gene and associated promoter.

4.2.2 DNA extraction and Tn-seq

For DNA extraction, a pellet of Mtb was obtained in an O-ring tube and supernatant removed. The pellet was then heated to 85°C for 20-30 minutes. After heat killing, the pellet was reimmersed in 0.6mL CTAB extraction solution (100mL water, 5.84g NaCl, 1g Hexadecyltrimethylammonium bromide, 0.2mL 0.5M EDTA). The sample was then transferred to a 2mL tube containing 1g 0.1mm zirconia beads and 0.6mL chloroform. The sample was then bead-beaten for 30 second at 7200rpm on a Presellys Evolution. The tube was centrifuged at 16,000g for 2 minutes. ∼0.4mL aqueous (top) phase was then pipetted into a fresh tube avoiding the white interface. 2 volumes of 100% ethanol was then added and the sample was mixed by inversion to precipitate DNA. In cases where DNA was visible by eye a 2 minute centrifugation was done, otherwise centrifugation was extended to 15 minutes. Supernatant was removed and DNA pellet was washed with 700uL 70% ethanol. The sample was briefly centrifuged again and ethanol was removed. Pellet was allowed to dry for 15 minutes before reimmersion in 100uL Tris-Cl. A nanodrop was used to measure DNA concentration.

4.2.3 Drugs

The drugs INH and RMP were used in these screens as they are part of the first line regimen used to treat Mtb. The other first line drugs (EMB and pyrazinamide) lack killing activity *in vitro* against Mtb and are only used for the first 2 months in the first line TB regimen. Ranges of drug concentrations were chosen based on estimates of the maximum concentration observed in human plasma and then performing dilutions to simulate reduced levels of drug. C_{max} was assumed to be 1.0 ug/mL for INH and 4.0 ug/mL RMP based on available data [20]. Drugs were dissolved in water and filter sterilized before use.

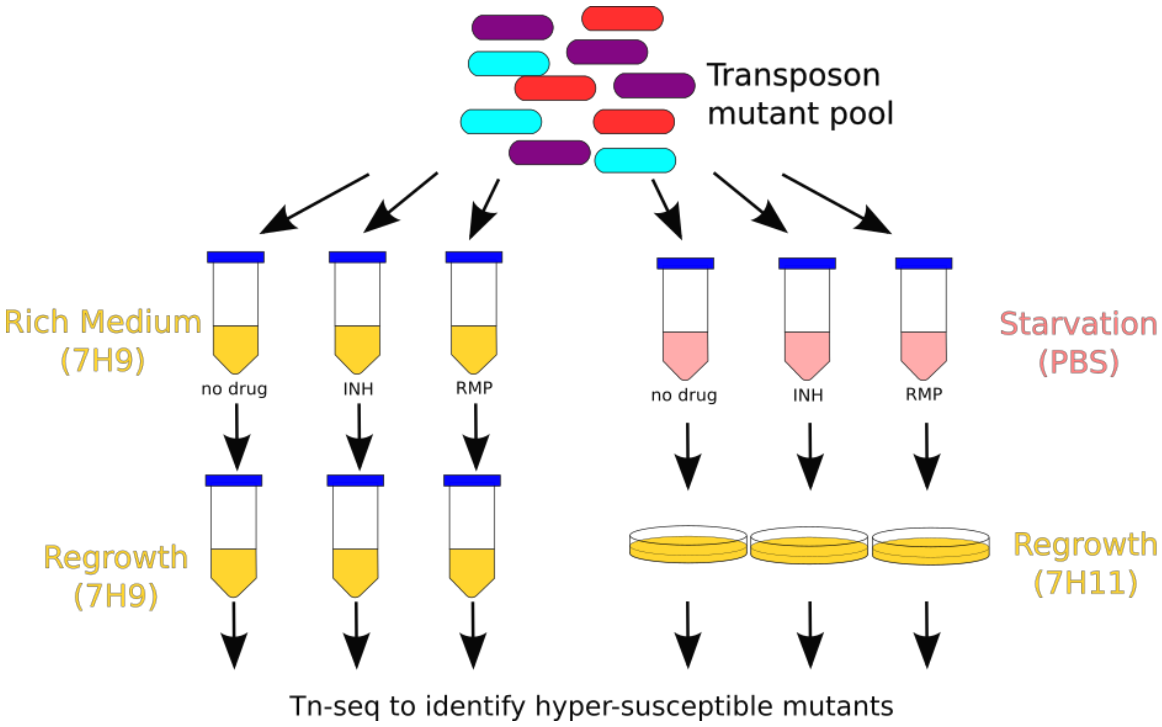

Figure 4.1: Schematic of experimental setup for identifying adjunctive drugs. A transposon mutant pool constructed in H37Rv is grown in a single vessel and aliquoted to tubes. After a period of adjustment to the new conditions, antibiotics (INH,RMP) are added.

4.2.4 Conditions for antibiotic hyper-susceptibility screens

Screen in nutrient-rich medium. The setup for our genome-wide differential susceptibility screens is sketched in figure 3.1. For screening in rich medium we inoculated a 1mL aliquot of the combined transposon mutant pool (consisting of approximately 1.2×10^6 unique mutants) into 200mL 7H9/Tw+/G-/OADC contained in a 1.3L roller bottle. This was shaken at 37°C for 3 days to reduce bacterial clumping. These samples were then split into 4 50mL aliquots and centrifuged (2000g for 5 minutes). Pellets were then reimmersed in 5mL 7H9/Tw+/G-/OADC and pooled into a single tube. The culture was then strained through a 40μm cell strainer to remove very large clumps. The OD was then taken and the culture diluted to OD 0.1. Diluted culture was aliquoted to 21 tubes (10mL/tube) for the screen. After 1 day of shaking at 37°C, drugs (dissolved in water) were the added to each tube (day 0). Final concentrations of INH were 0.01ug/mL, 0.1ug/mL, and 1.0ug/mL while final concentrations of RMP were 0.04ug/mL, 0.4ug/mL, and 4.0ug/mL. 3 replicates for each drug-concentration were performed as well as a no drug control (21 tubes total).

Cultures were checked daily for increased growth. If growth was observed to be near OD 0.8 then 800uL of culture was removed for OD measurement. If OD exceed 0.8 the culture was diluted 1:20 into fresh 7H9/Tw+/G-/10% OADC (with or without drug as appropriate). This passaging routine was intended to keep the cultures in log-phase throughout the screen. Each day (after dilution, if necessary) 400uL of culture were removed and processed for CFU enumeration. Briefly, samples were centrifuged (3000g for 5 minutes) and washed twice with 7H9/Tw+/G-/10%OADC to remove antibiotic. 10-fold dilutions were done in 7H9/Tw+/G-/10%OADC and 100uL of each dilution was plated on 7H11 agar. T-shaped spreaders were used to spread liquid evenly across plates. CFUs were counted after 34-36 days.

After 6 days of drug exposure, tubes were thrice centrifuged (2000g for 5 min) and washed with 10mL 7H9/Tw+/G-/OADC to remove drug before reimmersing in a final volume of 10.8mL. 0.8mL of each culture was then removed to confirm OD was below 0.4 and diluted appropriately otherwise. Tubes were passaged and diluted (1:20) in fresh 7H9/Tw+/G-/10%OADC upon reaching above OD 0.8. This was repeated twice (approximately 9-10 doublings). Samples were then centrifuged and processed for DNA extraction.

DNA library preparation was performed according to an earlier version of the protocol described in appendix B. Differences from the appendix protocol included use of Thermo-Fisher Taq polymerase with Thermopol buffer in place of NEB Q5 master mix (annealing temperatures of 58°C for both PCRs). Additionally there were some minor volume changes. Libraries were sequenced (2×100bp) on an Illumina HiSeq 2500 by Tom Ioerger and Aashish Srivastava of Texas A&M University. A total of 18 samples were sequenced yielding between 3,732,882 – 6,614,078 reads per sample for a total of 91,626,354 paired-end reads.

Screen in starvation medium. For screening the Mtb transposon mutant library in starvation conditions we first inoculated a 1mL aliquot of the combined transposon mutant pool into 200mL 7H9/Tw+/G-/10%OADC contained in a 1.3L roller bottle. This was shaken at 37°C for 1 day to reduce bacterial clumping. 50mL of this culture was then split into 5 10mL aliquots (in 50mL conicals) and incubated, shaking for 1 day. The OD was then taken and the cultures were pooled

and diluted to OD 0.01. Diluted culture was aliquoted to 60 tubes (10mL/tube). After cultures reached OD 0.85, cultures were repooled. We found growing the pools in these 10mL aliquots reduced bacterial clumping. After pooling, the culture was centrifuged (3000g for 5 min) and washed twice with PBS-Tw. The culture was then diluted to OD 0.52 in PBS-Tw. This was then aliquoted to 90 tubes (10mL/tube) to perform the screen. Immediately, 3 tubes were removed for CFU enumeration (day -14). After 14 days starvation in PBS-Tw an additional 3 tubes were removed for CFU enumeration and regrowth (day 0). Drugs (dissolved in water) were then added to the remaining tubes. Final concentrations of INH were 0.01ug/mL, 0.1ug/mL, and 1.0ug/mL while final concentrations of RMP were 0.04ug/mL, 0.4ug/mL, and 4.0ug/mL. Samples were taken for regrowth and CFU enumeration at day 7 and day 14. 6 replicates were performed for each drug-concentration and timepoint as well as no drug controls (84 tubes total). Only 3 samples from each group were prepped for sequencing to reduce sample costs.

CFUs were estimated for each sample by removing 400uL of bacterial culture, centrifuging (3000g for 5 minutes) and washing once with PBS-Tw (to remove antibiotic). 10-fold dilutions were done in PBS-Tw and 50uL of each dilution was plated on 7H11 agar. T-shaped spreaders were used to spread liquid evenly across plates. CFUs were counted after 25-35 days. For regrowth, the remainder of each tube was twice centrifuged (3000g for 5-10 min) and washed with 10mL PBS-Tw to remove drug. Centrifugation was done a final time and the sample was reimmersed in 250uL PBS-Tw. 50uL of washed transposon pool was plated on each of four 7H11 agar plates and spread with 10-15 3mm sterile glass beads. Samples were regrown until bacterial lawn formwed (7-14 days, depending on group). Lawns for the four agar plates were scraped and pooled into 2mL tubes for DNA extraction.

DNA library preparation was performed as described in appendix B. Libraries were sequenced (2×100bp) on an Illumina HiSeq 2500 by Tom Ioerger and Aashish Srivastava of Texas A&M University. A total of 36 samples were sequenced yielding between 1,508,879 – 6,085,006 reads per sample for a total of 134,634,944 paired-end reads.

Screen in hypoxic medium. An attempt to screen the Mtb transposon mutant library under hypoxic conditions was also made. Briefly, a shaker plate was placed inside an anaerobic chamber

utilizing a specialized gas mixture to remove oxygen from the chamber. Levels of oxygen was monitored using a chemical indicator to ensure hypoxia throughout the experiment. Cultures were setup in a similar fashion to the above screen in starvation medium except using 7H9/Tw+/G-/OADC in place of PBS-Tw. At OD 0.86, the culture was diluted to 0.54 and aliquoted to 84 tubes with vented caps to allow the anaerobic atmosphere to reach the liquid (we found that non-vented tubes were air-tight). All tubes were then transferred to the anaerobic chamber (day -7). After 1 day inside the chamber, stickers were added to block the vents in the cap in an attempt to prevent water loss. On day 0, (7 days after tubes entered anaerobic chamber) non-vented caps were added to improve water retention. After adding non-vented caps, antibiotics were added to tubes. Final concentrations of INH were 0.01ug/mL, 0.1ug/mL, and 1.0ug/mL while final concentrations of RMP were 0.04ug/mL, 0.4ug/mL, and 4.0ug/mL. Samples were taken for regrowth and CFU enumeration at day 7, day 14, day 21, and day 28. 3 replicates were performed for each drug-concentration and timepoint as well as no drug controls (84 tubes total). CFU enumeration and regrowth were done as written for our screen in starvation medium. CFUs were counted after 21-54 days. Due to issues encountered with low viability (see Results), these samples were not processed for sequencing.

4.3 Computational Analysis

LFC and p-values for each condition were calculated as described in section 3.3. For the rich medium screen, gene mutants were considered "differentially susceptible" to a drug if the absolute value of the $\log_2$-fold change (relative to no drug control) was greater than 0.5 and the adjusted p-value (Benjamini-Hochberg) was less than 0.05. For the starvation medium screen, gene mutants were considered "differentially susceptible" to a drug if the absolute value of the $\log_2$-fold change (relative to no drug control) was greater than 0.5 and the adjusted p-value was less than 0.05 at both the 7 days and 14 day time points.

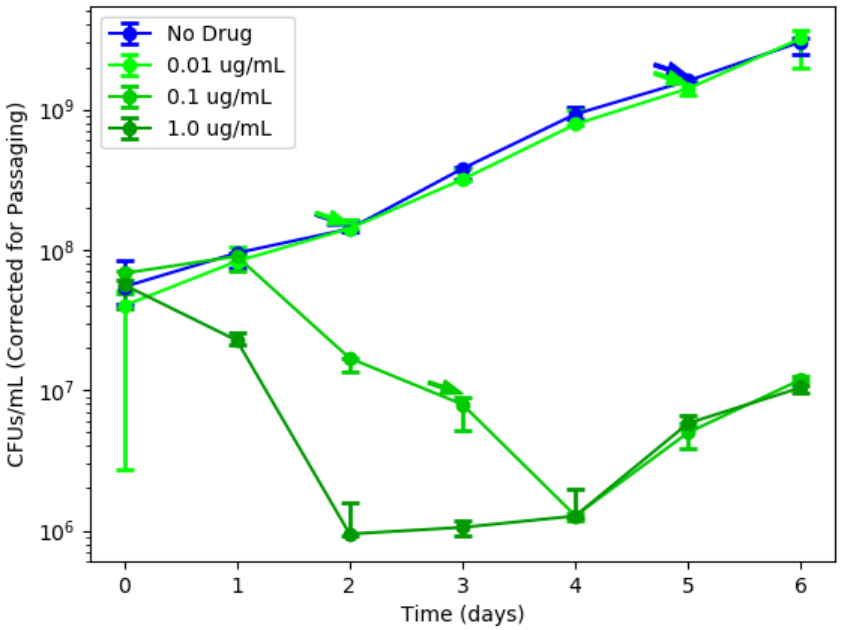

Figure 4.2: Bacterial viability of transposon mutant pools after addition of INH at 0 days. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

4.4 Results

4.4.1 Transformation efficiency

Upon transformation, we estimated each of our six independent H37Rv transposon mutant libraries contained between 1.7˜2.5 $\times$ 10^5 unique insertion events for a combined total of 1.2 $\times$ 10^6 unique events with ∼0.2% background. For all experiments we utilized a combined pool of all 6 independent pools.

4.4.2 Bacterial viability during screen

During each stress and antibiotic exposure, bacterial viability of the entire transposon mutant pool was monitored. Bacterial viability during rich medium is provided in figures 4.2-4.3. Viability during nutrient starvation is provided in figures 4.4 - 4.5. Viability during hypoxic conditions is provided in figures 4.6 - 4.7. Notably there was a large reduction in viability during exposure to hypoxia which prevented us from have enough material to proceed with Tn-seq for these samples.

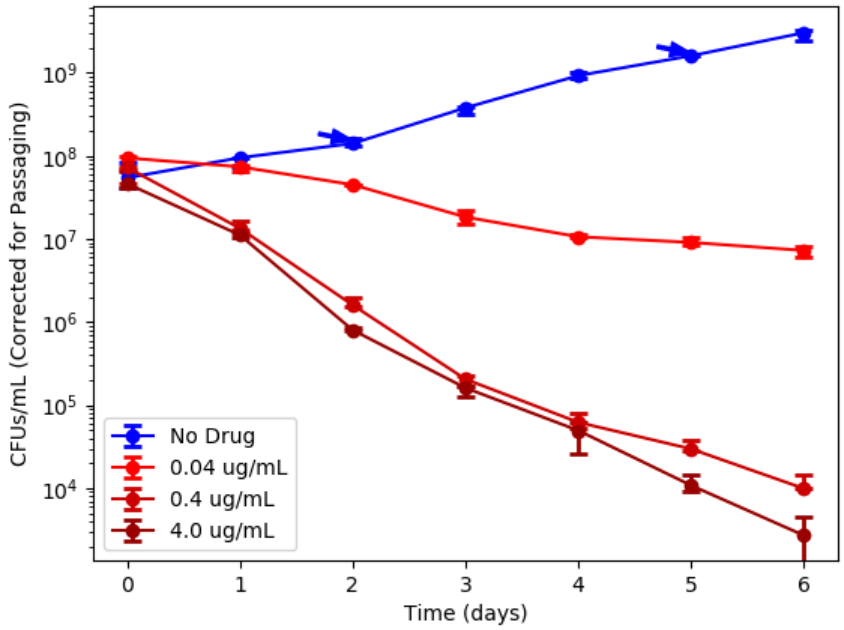

Figure 4.3: Bacterial viability of transposon mutant pools after addition of RMP at 0 days. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Points represent median. Error bars are minimum and maximum.

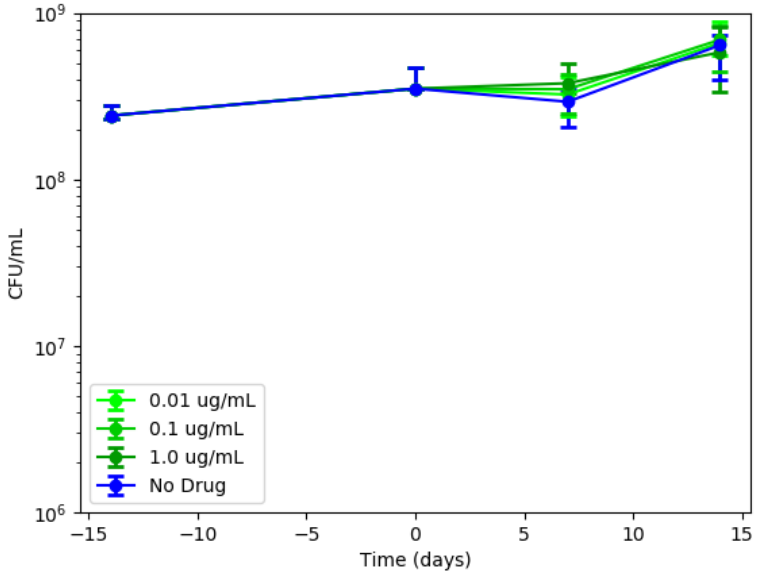

Figure 4.4: Bacterial viability of transposon mutant pools in PBS after addition of INH at 0 days. Final concentration of drug in each tube provided in legend. Between 3 and 6 replicates were collected at each time point (3 for day -14 and day -7). One outlier at day 7 were removed (likely a mislabeled sample) and one outlier at day 14 (cause unknown). Points represent median. Error bars are minimum and maximum.

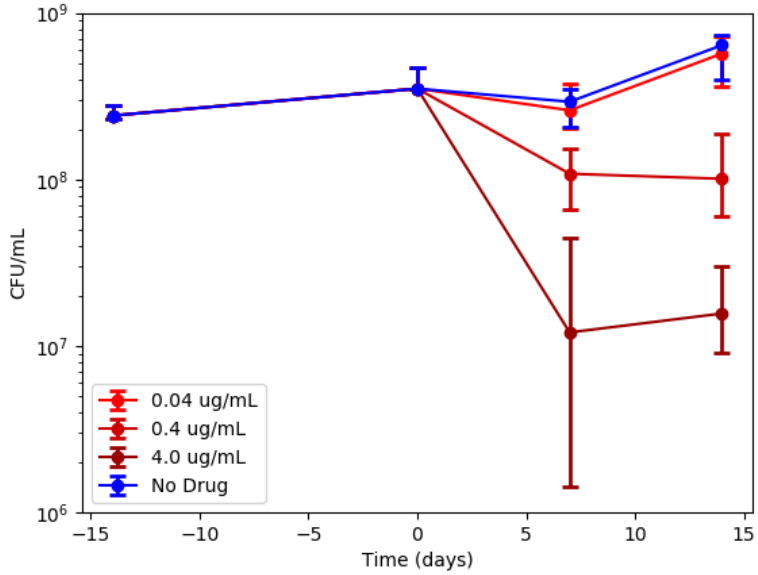

Figure 4.5: Bacterial viability of transposon mutant pools in PBS after addition of RMP at 0 days. Final concentration of drug in each tube provided in legend. Between 3 and 6 replicates were collected at each time point (3 for day -14 and day -7). One outlier at day 7 were removed (likely a mislabeled sample) and one outlier at day 14 (cause unknown). Points represent median. Error bars are minimum and maximum.

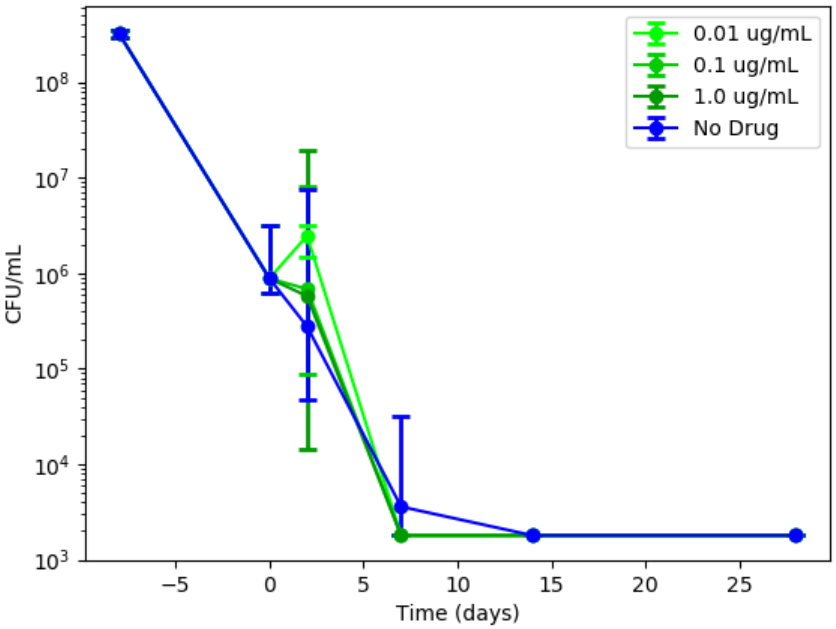

Figure 4.6: Bacterial viability of transposon mutant pools under hypoxia after addition of INH at 0 days. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Limit of detection at all timepoints was 2×10^3 CFUs/mL.

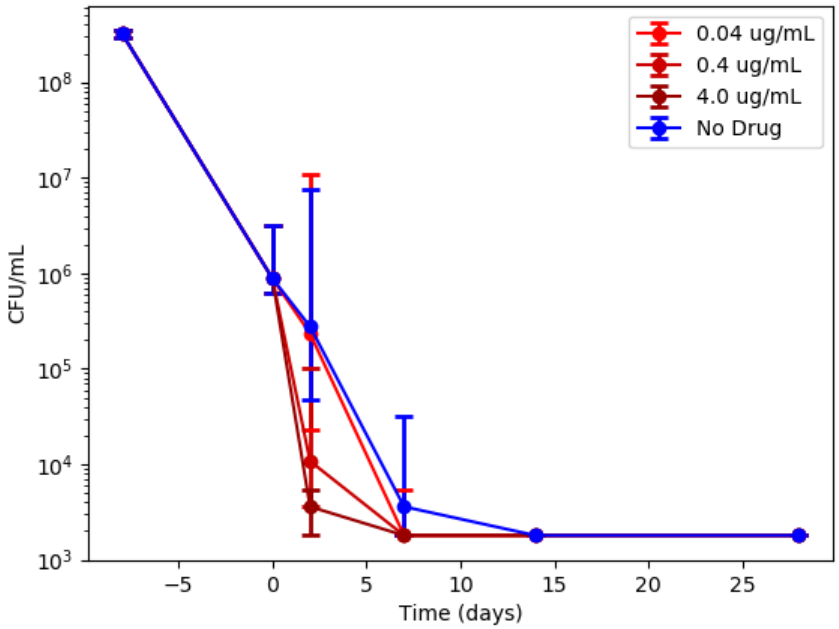

Figure 4.7: Bacterial viability of transposon mutant pools under hypoxia after addition of RMP at 0 days. Final concentration of drug in each tube provided in legend. 3 replicates were collected at each time point. Limit of detection at all timepoints was 2×10^3 CFUs/mL.

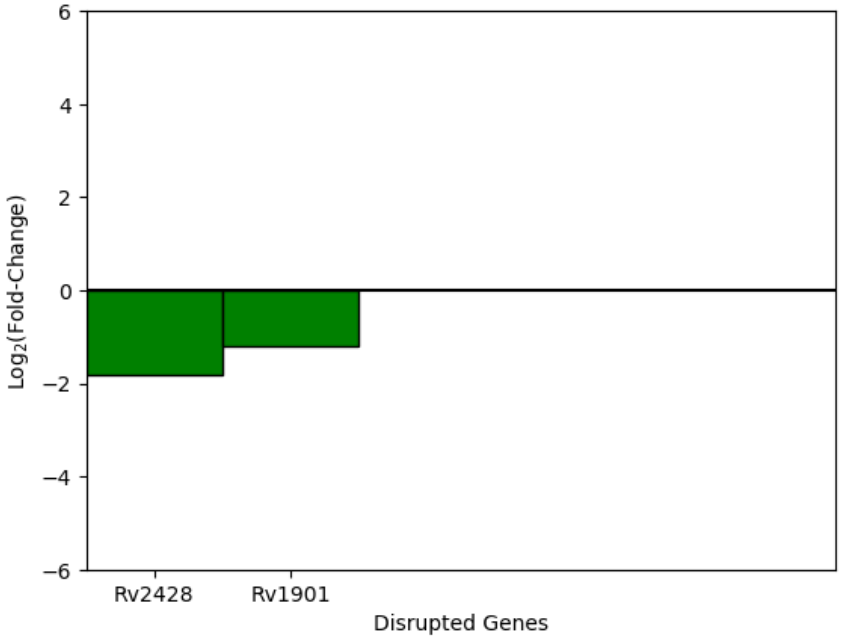

Figure 4.8: Effect size of disruption of genes in H37Rv (fold-change is normalized to no drug control) in rich medium (7H9). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to INH, positive values are hyper-tolerant ones.

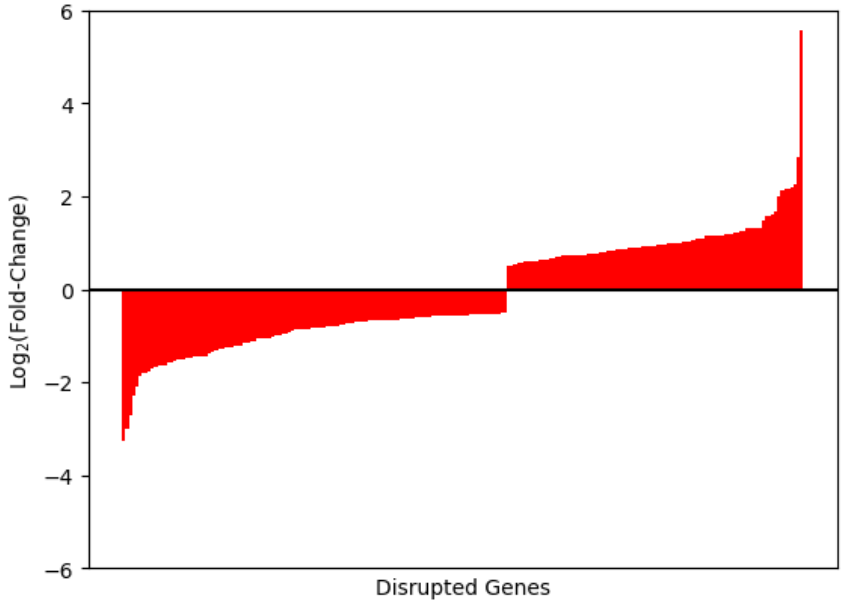

Figure 4.9: Effect size of disruption of genes in H37Rv (fold-change is normalized to no drug control) in rich medium (7H9). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to RMP, positive values are hyper-tolerant ones.

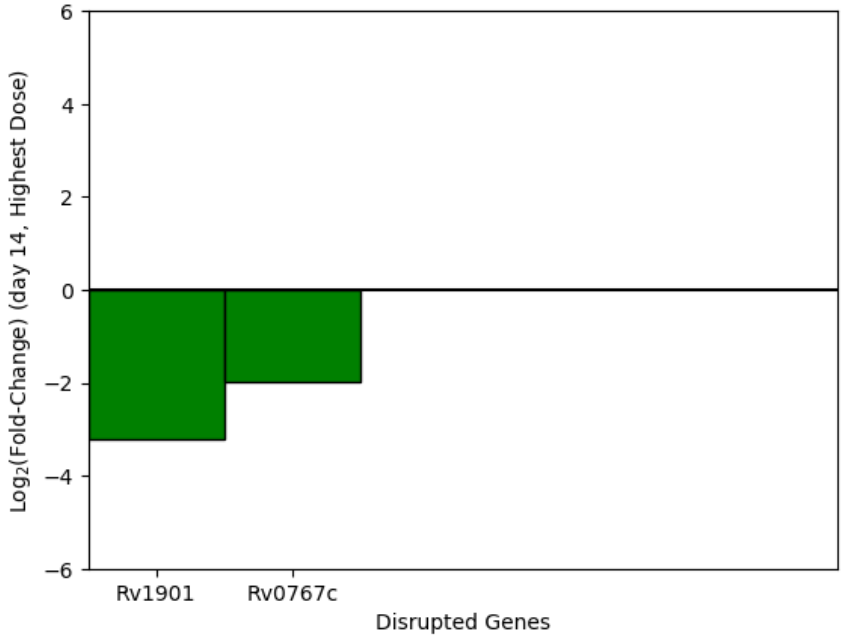

Figure 4.10: Effect size of disruption of genes in H37Rv (fold-change is normalized to no drug control) in nutrient starvation (PBS). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to INH, positive values are hyper-tolerant ones.

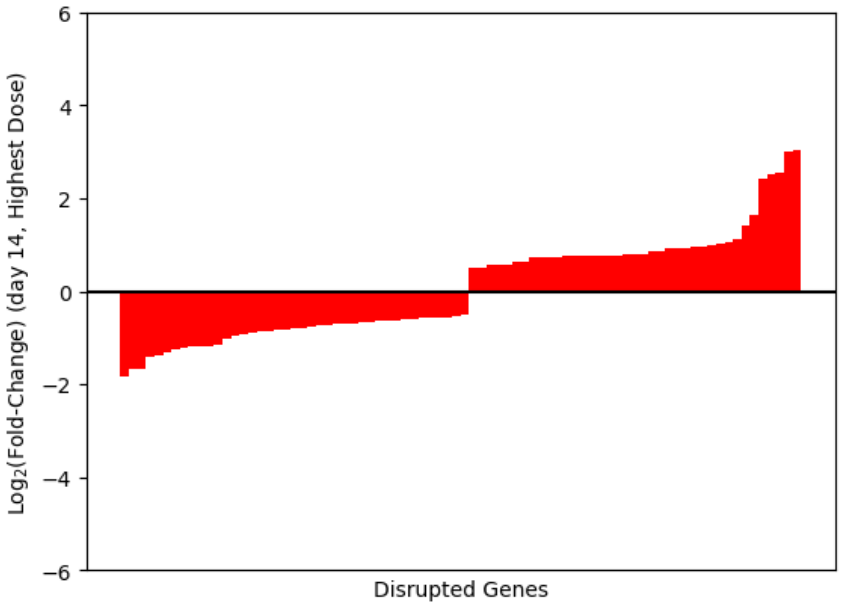

Figure 4.11: Effect size of disruption of genes in H37Rv (fold-change is normalized to no drug control) in Nutrient Starvation (PBS). Transposon disruption mutants are sorted from greatest effect size to smallest. Only genes for which effect size was statistically significant are plotted. Negative values represent mutants hyper-susceptible to INH, positive values are hyper-tolerant ones.

4.4.3 Mutants with environment-induced hypersusceptibility

In rich medium we identified 2 mutants as differentially susceptible to INH (2 hypersusceptible), 214 to RMP (121 hypersusceptible). In starvation medium we identified 2 mutants as differentially susceptible to INH (2 hypersusceptible), 80 to RMP (41 hypersusceptible). Significant mutants and effect sizes are provided in appendix D. Effect sizes for these mutants are plotted in figures 4.8 - 4.11.

4.5 Validation and exploration of a hypothesized mechanism

4.5.1 Antibiotic susceptibility of the caeA mutant

Our results showed that in rich medium the caeA (Rv2224c) transposon mutants were hypersusceptible to RMP ($\log_2$ fold-change = -1.8). Therefore we decided to validate our results using a knockout mutant in caeA (caeA-KO) and a mutant strain where the knocked-out gene is complemented with a second copy (caeA-COM). We tested the antibiotic susceptibility of the caeA-KO, caeA-COM and background CDC1551 strains. The results are reported in table 4.1. Notably, a few drugs showed

Table 4.1: MIC of wildtype, caeA-KO, and caeA-COM strains to various antibiotics as measured by Alarmar blue assay. All MICs were measured in triplicate.

	CDC1551-WT	**caeA-KO**	**caeA-COM**
Clarithromycin	4.0625	1.01563	4.0625
Rifampin	0.04	0.01	0.04
Isoniazid	0.03	0.03	0.03
Ethionamide	5-10	10	10
Meropenem	7.13	3.56-7.13	7.13
Amikacin	0.25	0.125	0.25
Moxifloxacin	0.05	0.05	0.05
Vancomycin	20	2.5	20
Ethambutol	1	0.5	1
Chloramphenicol	5	2.5	5
Desacetyl-rifampin	0.038	0.0095	0.038

changes in MIC of 4× or more including rifampin. This provides some validation that this mutant is hypersusceptible to this drug.

4.5.2 De-acetylation as a mechanism of antibiotic tolerance

After we observed that the caeA gene appears to be involved in tolerance to RMP we speculated on roles for the gene. Notably, CaeA has been shown to have carboxyesterase activity[70]. Given that RMP is an ester we speculated that CaeA might remove the ester group from RMP. Furthermore, crystallographic data from the target of RMP (RNA polymerase) shows that the RMP ester group is likely involved in binding to its target [16]. This suggested a mechanism for our observed hypersusceptibility phenotype: In wildtype bacteria CaeA cleaves the ester group from RMP, which reduces binding affinity to RNA polymerase and reduces the effectiveness of the drug. When CaeA is absent RMP is not cleaved and rifampin activity appears to increase due to increased binding affinity.

To test this hypothetical tolerance mechanism we compared the activity of desacetyl-rifampin to rifampin in the wildtype and caeA-KO strain. If our hypothesis was correct then the caeA-KO should not be hypersusceptible to desacetyl-rifampin (relative to the wildtype). The last row of table 4.1 shows that the caeA-KO strain is also hypersusceptible to desacetyl-rifampin - proving our hypothesis incorrect. In conclusion the mechanism by which the caeA-KO mutant is hypersusceptible to rifampin is not mediated through deacetylation.

Table 4.2: Mutants (Mtb) differentially susceptible to rifampin in both rich medium (7H9) and starvation medium (PBS). Numbers represent log fold change between rifampin-treated group and no-drug control. A negative LFC indicates a mutant is hypersusceptible to the drug while a positive LFC indicates it is hypertolerant. Samples in rich medium were treated with 0.04ug/mL rifampin. Sample in starvation medium were treated with 4.0ug/mL. Only significant mutants with an effect size greater than 1 in either direction ($|LFC| > 1$) in both conditions are included.

Gene	**7H9 6d LFC**	**PBS 7d LFC**	**PBS 14d LFC**	**Annotation**
Rv0049	-2.99	-1.45	-1.81	hypothetical protein
Rv0199	1.14	4.5	2.53	membrane protein
Rv0200	1.61	4.69	2.56	transmembrane protein
Rv0655	1.32	5.03	3.01	ABC transporter ATP-binding protein
Rv0819	1.33	-1.44	-1.67	mycothiol acetyltransferase
Rv0994	-3.25	-1.27	-1.01	molybdopterin molybdenumtransferase 1
Rv2179c	-2.72	-1.32	-1.3	3'-5' exoribonuclease
Rv2690c	5.56	3.26	1.63	integral membrane protein
Rv2709	1.16	-1.26	-1.16	transmembrane protein
Rv3723	1.18	4.94	3.04	transmembrane protein

4.5.3 Role of environment in hypersusceptibility

To explore the role the environment plays in hypersusceptibility to antibiotics we compared significant genes from the rich medium and nutrient starvation screens. Table 4.2 presents mutants with differential susceptibility to rifampin in both rich and starvation medium. Table 4.3 presents mutants with differential susceptibility to rifampin in starvation medium, but with an effect size (LFC) near zero or in the opposite direction in rich medium. Thus, the differential susceptibility phenotype of these mutants appears to be significantly impacted by the environment.

4.6 Summary and discussion

We identified hundreds of Mtb transposon mutants with survival phenotypes relative to wildtype bacteria. Nearly all of these were specific to RMP with only 2 mutants each detected as hypersusceptible to INH in nutrient starvation or rich medium. There are a few possible explanations for the observed differences in the number of detected phenotypes between the drugs. First, in the case of the nutrient starvation model (PBS) 3 different RMP concentrations were used for hypothesis testing. This greatly improves the statistical power for detection using the JT test. Secondly, in the rich medium model the dose of INH may have been too low - as evidenced by no difference in the CFU counts for the bulk culture between the low-dose INH group and the no drug control (4.6).

Table 4.3: Mutants (Mtb) which may be involved in starvation-induced tolerance to RMP. Numbers represent log fold change between rifampin-treated group and no-drug control. A negative LFC indicates a mutant is hypersusceptible to the drug while a positive LFC indicates it is hypertolerant. Samples in rich medium (7H9) were treated with 0.04ug/mL rifampin. Samples in starvation medium (PBS) were treated with 4.0ug/mL. To compute this list, statistically significant mutants with an effect size greater than 1 in either direction ($|LFC| > 1$) in starvation medium were first identified. These identified mutants were then examined in data collected from rich medium. Mutants included in the list below are significant in starvation medium and the product of the 7H9 LFC and PBS 14d LFC is less than 0.5 (suggesting an effect near zero or in the opposite direction in rich medium).

Gene	**7H9 6d LFC**	**PBS 7d LFC**	**PBS 14d LFC**	**Annotation**
Rv0458	-0.11	2.3	1.05	aldehyde dehydrogenase
Rv0819	1.33	-1.44	-1.67	mycothiol acetyltransferase
Rv0989c	0.03	2.77	1.01	polyprenyl-diphosphate synthase GrcC
Rv0998	-0.22	-1.24	-1.25	acetyltransferase Pat
Rv1183	0.1	0.52	1.12	transmembrane transport protein MmpL10
Rv1908c	0.85	-0.88	-1.15	catalase-peroxidase
Rv2051c	-0.19	-1.33	-1.23	polyprenol-monophosphomannose synthase
Rv2199c	-0.5	2.53	2.42	cytochrome c oxidase polypeptide 4
Rv2374c	-0.2	-1.47	-1.68	heat-inducible transcription repressor HrcA
Rv2392	0.95	-1.29	-1.17	phosphoadenosine phosphosulfate reductase
Rv2633c	0.41	1.76	1.02	hypothetical protein
Rv2709	1.16	-1.26	-1.16	transmembrane protein
Rv2733c	0.01	-1.55	-1.19	(dimethylallyl)adenosine tRNA methylthio-transferase
Rv3680	-0.26	-1.12	-1.4	anion transporter ATPase
Rv3923c	-0.17	-1.1	-1.36	ribonuclease P protein component

This may have reduced the effect size of mutants hypersuseptible to INH. For future work, it appears that bacterial viability is highly sensitive to INH as a 10-fold higher dose produced too much killing to be useful for Tn-seq. Therefore, future attempts might consider either using INH doses between these two concentrations or else reducing the duration before regrowing the bacteria to ensure lower bacterial killing at when using a higher antibiotic concentration. Lastly, it is of course possible that the number of genes involved in persistence during INH exposure is simply substantially less than for RMP.

The screens performed here also had a number of limitations. As mentioned above, the INH concentration may not have been sufficiently high to detect many hypersusceptible mutants in rich medium. Secondly, we utilized two different regrowth protocols for the rich medium and the starvation conditions. This occurred as a result of the decision to improve our protocol when processing the starvation samples (which were collected months after the rich medium samples). This choice greatly simplified and accelerated processing primarily by using solid medium to regrow the pools after antibiotic exposure instead of regrowing in liquid medium. While the components of the liquid medium (7H9) are largely similar to those of the solid medium (7H11) the liquid medium included Tween-80 (to avoid clumping) and calcium chloride while the solid medium contained agar, casein hydrolysate, and glycerol. However, we do not anticipate these differences to make a substantial different in identified mutants as independent no-drug controls generated in identical conditions were utilized for all comparisons.

As evidenced in figures 4.6 and 4.7 our attempt to establish a hypoxia model for identifying hypersusceptible mutants was a failure due to rapid and profound bacterial viability loss. Notably our results conflict with those previously published for hypoxia models [105, 95] which did not observe such severe viability loss in similar models. There were some differences between these models and ours including that our oxygen removal was especially rapid due to our use of an anaerobic environment, vented tubes, shaking, and a high viability culture - all of which will contribute to a rapid decline in oxygen concentration in the culture. It is possible that the bacteria are unable to adjust quickly to the oxygen-free environment and thus are killed quickly in our model. Another

possibility is that the exclusion of glycerol from our medium may serve some unknown protective effect.

Interestingly in the nutrient starvation model the observed effect sizes for hypersusceptibility were somewhat small after 14 days of rifampin treatment (greatest effect size was -1.81, $\log_2$). This suggests that persistence to rifampin under nutrient starvation may involve multiple genes each with a small contribution to survival or, similarly, genes with redundant function. An interesting follow-up experiment would be to generate combined knockouts in multiple genes to investigate if gene disruptions are additive or synergistic. Synergism would suggest redundancy while additivity would suggest multiple independent pathways for survival are at work. Also, the bacterial viability in our nutrient starvation model treated with RMP was notably not reduced between day 7 and day 14 which is somewhat surprising given the bacterial viability reduction observed between day 0 and day 7. There are at least two possible explanations. First, it may be that there is only a small population of bacteria that are susceptible to RMP under nutrient starvation. These are all killed between day 0 and day 7 and therefore no susceptible population remains at day 7. Secondly, some experimental results suggests that rifampin can quickly degrade *in vitro*. Therefore it is possible that rifampin is degraded in our model before day 7.

We decided to validate the caeA mutant due both to its clear hypersusceptible phenotype in nutrient rich medium and the convenient availability of the knockout and complemented strain. Our results clearly showed hypersusceptibility in a growth based assay. Future work should also test for hypersusceptibility in a killing-based assay (ie a decrease in CFU) during exposure to rifampin, which may be most relevant for optimizing bacterial killing and understanding persistence. Interestingly, the caeA mutant did not show hypersusceptibility to other drugs including INH and MOX. This suggests that survival in the presence these other drugs may involved entirely independent pathways. Thus persistence may be a drug-specific phenomenon. Lastly, our results clearly indicated that the mechanism by which the caeA mutant is hypersusceptible to RMP is not through deacetylation. Future work should be carried our the identify the true mechanism which may help to optimize RMP-containing regimens for TB and, hopefully, decrease therapy duration.

Chapter 5

Discussion and conclusions

5.1 The essential genes of Mav

We identified 230 genes as essential in both Mav and Mtb (appendix E). These may represent particularly good targets for drug development, as inhibitors of a gene product are likely to be effective against a close ortholog. As expected, a number of well-demonstrated targets are present. This includes the targets of the mycobacterial drugs cycloserine (alanine racemase, D-alanine – D-alanine ligase), rifamycins (RNA polymerase beta subunit), macrolides (50S ribosome), aminoglycosides (30S ribosome), fluoroquinolones (type IV topoisomerases and gyrases), bedaquiline (ATP synthase), and ethambutol (arabinosyltransferase). Additional compounds that have been reported to have some activity against mycobacteria include tryptophan synthase inhibitors [2], ClpP inhibitors [21], and Rho inhibitors (albeit only shown to be effective through genetic manipulation) [12]. A brief literature search also reveals many compounds that inhibit non-mycobacterial orthologs of these genes but appear to lack published results for killing activity in mycobacteria including inhibitors of GroEL [61], RibBA [52], SecA [54], and LigA [14, 38]. It is thus apparent that many opportunities are available for targeting these overlapping essential mycobacterial genes.

Our analysis classified four protein-coding genes on the two plasmids as essential (3 on pMAC109a and 1 on pMAC109b). This was somewhat surprising, as these plasmids have multiple copies per cell, and a disruption of a single gene copy should, in theory, be complemented by other copies. We used NCBI BLAST to find homologs of these genes. DFS55_24645 (on pMAC109a) and DFS55_25425 (on pMAC109b) are homologous to Rep, a protein critical for the replications of plasmids. Thus,

one possible explanation for the essentiality of these Rep homologs is that plasmid copy number will decrease in daughter cells inheriting the plasmid (with no plasmid replication possible in a cell with all copies containing disrupted Rep). This is a strong selective pressure against the mutant plasmid. DFS55_14680 (on pMAC109a) is a ParA homolog. ParA controls the distribution of plasmids to daughter cells such that cells inherit the plasmid more equally. It is not immediately apparent how a more random distribution of the plasmids due to disruption of ParA would lead to a growth defect. Lastly, DFS55_24600 (on pMAC109a) is a hypothetical protein also classified as essential. It lacks a paralog in the MAC109 chromosome and an ortholog in Mav strain 104 (which does not contain plasmids). Thus, it appears to be non-essential for the Mav subsp. *hominissuis* pangenome. DFS55_24600 is homologous to Rv3081 from Mtb and our analysis identified Rv3081 as "GD" (approximately 0.25 Relative Fitness). It is also apparent from examining the raw Tn-seq read counts that transposon insertion in the beginning of this gene does not have a profound effect on growth rate in MAC109 (this trend is less clear in H37Rv). Given these observations we can only speculate that this gene is addictive in MAC109 (and weakly addictive in H37Rv) – and may represent a toxin-antitoxin fusion with the toxin domain near the N-terminus. Future work could clone DFS55_24600 into an episomal (non-integrating) mycobacterial shuttle vector (such as pPB10) and examine the retention of the episome with and without this gene in the absence of antibiotic selection. Additionally, an attempt could be made to isolate a MAC109 mutant cured of pMAC109a.

Our analysis method to detect essential genes has several advantages over other methods, including its high robustness as the result of using the zero-inflated negative-binomial to model read counts, which can more accurately account for non-saturating libraries, as these have a high probability of a site having no observed insertions. This may be especially important for transposons which cannot easily achieve saturation without very large numbers of transformants (e.g. due to lack of strict TA site bias of Himar1), such as the Tn5 system [67]. Also, we have fully exploited the statistical independence of samples, which increases our statistical power. Other models, such as hidden Markov models, generally pool samples, limiting the usefulness of having biological replicates. However,

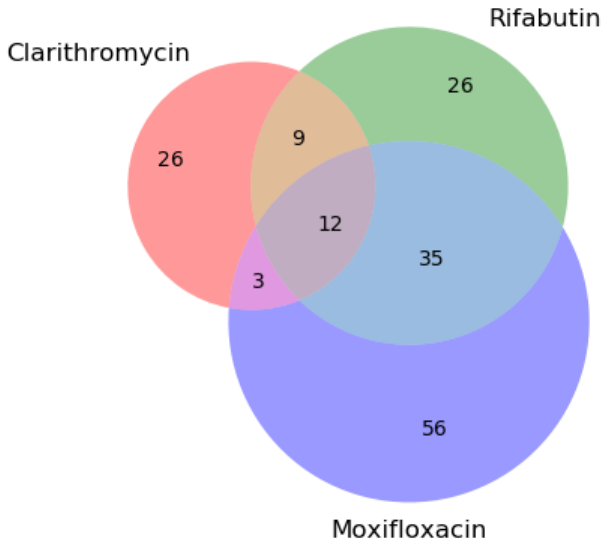

Figure 5.1: Venn diagram of Mav genes that, when disrupted, lead to hypersusceptibility to multiple drugs.

our method also has limitations. Using our collected data, we detected a somewhat low number of essential features in MAC109 relative to H37Rv (270 and 738, respectively) despite evidence that the genome was saturated with insertions (Figure 1). Most likely, this is due to our somewhat low sample size (5 independent libraries). Therefore, we believe that sequencing additional independent transposon mutant libraries could significantly increase the statistical power to detect essential genes in MAC109, particularly for features with fewer insertion sites. A previous study [24] used 14 independent libraries for H37Rv, which seemed to give our method good statistical power and may be a useful sample-size target for future studies. Additionally, while our method can correctly handle sites with low rates of insertion (e.g., [CG]GNTANC[CG]) it is possible that additional such sites exist that have not yet been defined. Defining the sites with low rates of insertion is especially important to avoid features falsely classified as essential.

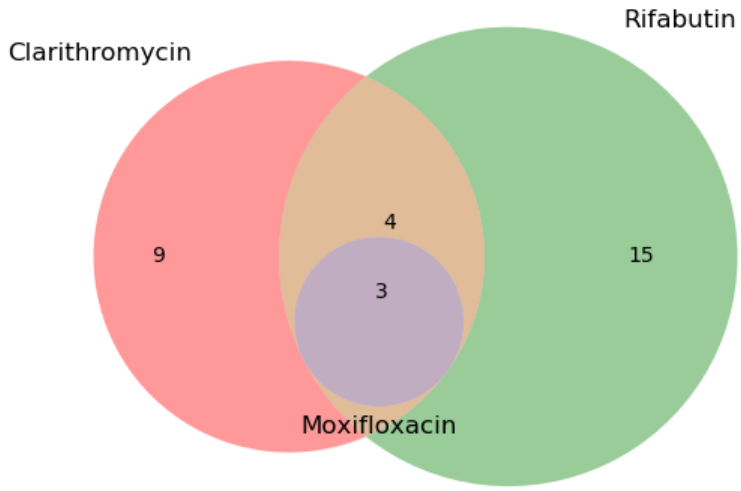

Figure 5.2: Venn diagram of Mav genes that, when disrupted, lead to hypertolerance to multiple drugs.

5.2 How specific are hypersusceptible mutants to a drug?

It was hypothesized that there might be a single pathway that controls the persister phenotype in the presence of multiple drugs. However, due to the use of a single set of samples as the control group there will be some expected positive correlation between groups. This complicates using the high-throughput data to assess this question as some overlap is expected due to sharing of the controls. Noting this limitation, a comparison of hypersusceptible mutants between drug classes shows that most of the mutants hypersusceptible to CLR and MOX were unique for those drugs 5.1. About 30% of mutant hypersusceptible to RFB were unique to this drug. These results suggest that many mutants are hypersusceptible only to a specific drug as opposed to all antibiotics classes. Furthermore our data shows that in the specific cases of the DFS55_12665 mutant in Mav and the caeA/Rv2224c mutant in Mtb that hypersusceptibility was somewhat specific to particular drug classes. The DFS55_12665 mutant was clearly hypersusceptible to MOX but lacked any detectable hypersusceptibility to CLR. These results suggests that there may not be a single antibiotic survival

pathway or "persistence pathway" but rather a persistence pathway specific for each drug of those we tested. It is possible that elements of these pathways may overlap which would explain the hypersusceptible mutants shared between drugs.

Similarly it was also hypothesized that there exists a single death pathway that is triggered by all antibiotic classes. The overlap of a small set of genes involved in death due to MOX,CLR, and RFB is intriguing 5.2. This hints at an overlap in at least part of the pathway, though the issue of bias due to shared controls as discussed above is still a risk. Perhaps most interesting is pyruvate kinase (DFS55_10765) which was found to confer hypertolerance to MOX, CLR, and RFB in Mav. Pyruvate kinase converts phosphoenolpyruvate and ADP to pyruvate and ATP, which occurs irreversibly under physiological conditions[88]. Thus, removal of this enzyme should lead to a decrease in the concentration of pyruvate. Interestingly, it was recently shown that accumulation of pyruvate can lead to increased killing by moxifloxacin [30]. Thus reduced pyruvate levels by removal of pyruvate kinase might be expected to produce the opposite effect. However, this effect was shown to result from the activity of pyruvate oxidase for which there is no annotated homolog in *Mycobacteria*. To explain this, there may be a weakly homologous enzyme with pyruvate oxidase activity in *Mycobacteria*. Alternatively, hypertolerance to antibiotics caused by increased pyruvate levels might be mediated through another mechanism entirely. Additional work should confirm the role of pyruvate kinase in antibiotic death after exposure to MOX, CLR, and RFB using isolated mutant strains. After validation, work should examine the accumulation of phosphoenolpyruvate (and also reactive oxygen species, which are produced by pyruvate oxidase) in this mutant before and during early exposure to these antibiotics.

5.3 Role of environment in hypersusceptibility

We found that mutations in the gene Rv1901/cinA cause hypersusceptibility to INH in both rich and starvation media with a particularly strong effect in starvation (appendix D). There is some weak evidence that Rv1901 may be a nicotinamide amidase (KEGG database: `https://www.genome.jp/`

dbget-bin/www_bget?mtu:Rv1901%5C%5C). Given INH has a similar molecular structure to nicotinamide it is possible that cinA degrades INH directly or intrabacterial products of INH, potentially reducing the toxicity of these products in the cell. We also identified mutants in Mtb which are differentially susceptible to rifampin in both rich medium and starvation medium in table 4.2. Notably, the mutants which were hypertolerant in PBS can be seen to have smaller LFCs at day 14 than at day 7. This suggests that these gene disruptions primarily delay the killing achieved by rifampin. Many of these are membrane proteins so one possible mechanism is that these membrane proteins, when present, allow rifampin into the cell. Thus removal of these channels will delay rifampin's entrance and killing activity. We also note that Rv2179c has been shown to be structurally closely related to RNase T[1], which is involved in processing various functional RNAs (mRNAs, tRNAs, etc). Therefore one possible mechanism for rifampin hypersusceptibility in these mutants is that bacteria with unprocessed functional RNA molecules are particularly impacted when rifampin binds the RNA polymerase - the only source of new RNA.

Interestingly, we identified two gene disruptions which lead to hyptertolerance in rich medium (7H9) but hypersusceptibility in starvation medium (PBS). These genes may be particularly interesting as these mutants seem to be less affected by the environmental changes (ie they have reduced killing in rich medium but increased killing in starvation medium). It may be that these genes are involved in the mechanism by which bacteria become hypertolerant during nutrient starvation. To further explore this we also identified mutants with large LFC in starvation medium but with small or opposite direction LFC in nutrient rich medium (table 4.3). We found that Rv2374c/HrcA causes strong hypersusceptibility to rifampin in PBS but not 7H9. Additionally we found that the mutants in the gene Rv1908c/katG confer hypersusceptibility to rifampin (but not 7H9). This may be due to katG preventing the accumulation of reactive oxygen species duing PBS exposure - which has been previously been linked to antibiotic activity in other bacterial species[102].

5.4 Adjunctive drug targets common to both Mav and Mtb

As we tested a rifamycin in both Mav and Mtb, we wondered whether there were overlaps in the genes involved in differential susceptibility between the two organisms - suggesting conserved mycobacterial pathways involved in rifamycin activity and potential targets for inhibition in multiple pathogens. We identified 8 gene disruption that cause strong hypersusceptibility and 1 gene disruption that causes strong hypertolerance ($|\text{LFC}| > 1$) in both Mtb and Mav.

Rv0820 (PhoT) is annotated as a phosphate ABC transporter. Notably, it is a close homolog of the active phosphate transporter PstB. Rv0929 and Rv0930 code for the PstA and PstC components of the PstA-C active bacterial phosphate transporter[50]. This suggests a connection between phosphate transport and rifamycin killing. It is possible that low intrabacterial phosphate levels may boost rifamycin activity in *Mycobacteria* though the rich medium is not phosphate limiting so this would need to be explored further. Rv2224c/caeA is a protease (and carboxylesterase) which we've validated in H37Rv to be hypersusceptible to rifampin 4. This represents a particularly attractive target to boost the activity of rifamycin-containing regimens in multiple *Mycobacteria* given its wide distribution and substantial effect size. Interestingly, an inhibitor is already available for this protein which could be further optimized to potentially reduce the duration of rifamycin-containing regimens [78]. Disruption of Rv2179c, which appears to cause hypersusceptibility to rifampin in nutrient starvation, may also trigger rifamycin hypersusceptibility in *Mycobacteria* generally. As speculated above, this may be related to its hypothesized role in processing functional RNA molecules. Additionally, proteins Rv0049, Rv1836c, and Rv3005c have no known function but were also found to have strong hypersusceptibility to rifamycins. The single gene disruption found to confer hypertolerance to rifamycins was Rv0819, annotated as a mycothiol acetyltransferase. This is somewhat surpising given that mycothiol synthesis has been linked to rifamycin hypersusceptibility [79]. Notably, the gene for Rv0819/DFS55_21365 is oriented such that a transposon disruption in this gene could affect the downstream transcription of the nearby gene Rv0820/DFS55_21345. It has been noted previously that the kanamycin promoter within the Himar1 transposon can serve as a promoter for downstream gene expression[86]. Therefore it is possible that a transposon insertion in this gene

Table 5.1: Mutants differential susceptible to rifamycins in Mtb and Mav. Significant mutants with an effect size greater than 1 in either direction (|LFC| > 1). Negative LFC implies the mutant is hypersusceptible whereas positive LFC implies it is hypertolerant to rifampin in rich medium.

Mtb Gene	**Mav Gene**	**Mtb LFC**	**Mav LFC**	**Mtb Annotation**	**Mav Annotation**
Rv0049	DFS55_00355	-2.99	-1.09	hypothetical protein	hypothetical protein
Rv0819	DFS55_21365	1.33	1.23	mycothiol acetyltransferase	mycothiol synthase
Rv0820	DFS55_21345	-1.42	-1.69	phosphate ABC transporter ATP-binding protein PhoT	phosphate ABC transporter ATP-binding protein
Rv0929	DFS55_20215	-1.23	-1.87	phosphate ABC transporter permease PstC	phosphate ABC transporter permease subunit PstC
Rv0930	DFS55_20210	-1.16	-1.0	phosphate ABC transporter permease PstA	phosphate ABC transporter permease PtsA
Rv1836c	DFS55_12730	-1.45	-1.44	hypothetical protein	hypothetical protein
Rv2179c	DFS55_14810	-2.72	-1.09	3'-5' exoribonuclease	hypothetical protein
Rv2224c	DFS55_15065	-1.8	-2.17	carboxylesterase A	alpha/beta hydrolase
Rv3005c	DFS55_07355	-1.25	-1.06	hypothetical protein	DoxX family protein

promotes survival during exposure to rifamycins as a result of increased expression of Rv0820 which, as noted above, may be involved in survival to rifamycins.

5.5 Limitations and future directions

One of the major limitations of this work is that essential genes cannot be assessed for their effect on hypersusceptibility. Likely this can be addressed in future work utilizing new high-throughput "knockdown" technologies which allow for reduced levels of particular genes to be achieved[85]. Much of the experimental setup and analysis developed here for our approach using the Tn-seq technology could easily be transferred to an effort utilizing these newer techniques for understanding persistence and antibiotic mechanisms.

Of primary importance for future work is the confirmation of the predictions from high-throughput data. As discussed in the introduction, these high-throughput techniques often have substantial noise and therefore must be validated, ideally with an entirely independent methodology. Here, this can likely most easily be accomplished by creation of knockout strains via a technique other than transposon mutagenesis. New methods allow for knockouts to be quickly generated in Mtb[76] and are likely functional in other *Mycobacteria* such as Mav. In designing validating assays a focus should

be placed on assessing increased killing - not just inhibition. Drugs such as ethambutol are known to have good inhibitory properties but have very little killing activity in Mtb. Here we have confirmed increased killing in Mav mutants using CFU counts, which are relatively robust but time consuming. Assays using bioluminescence or other ATP measurements may accelerate the ability to assess mutants for increased antibiotic killing [84]. However, a final validation should still be done with CFU enumeration or other direct viability assay (ie not ATP based).

In assessing multiple drugs, multiple environments, and multiple bacteria we have generated a wide array of potentially useful data for understanding antibiotic activity. However, it also became clear that validation studies for all the various conditions and experimental setups would be challenging. Therefore, future work should seek to focus on one particular drug and one bacterial species, potentially in a few easily achieved *in vitro* environments. This focus should allow for stronger and more precise conclusions to be made about antibiotic persistence.

Additionally, multiple properties of Mtb make experiments with this organism exceedingly difficult. These include its slow growth rate, preference for specialized medium (which is labor-intensive to make and generally cannot be ordered pre-made through biotech companies), the need for specialized facilities, and the propensity to form clumps in liquid culture. This last feature led to many failed experiments during this work. Mav grows significantly faster and is generally safer to handle though clumping is still a major issue - especially for interpreting CFU counts. Follow-up work to understand the mechanisms underlying the observed changes in antibiotic activities should first utilize more efficient model organisms such as *E. coli* and *M. smegmatis*. This approach will provide an opportunity to optimize assay conditions without wasting samples and time working with the more difficult model species. However, it is essential that discoveries in these simpler model organisms eventually be validated in the less-efficient, but more clinically-relevant ones.

Future work should utilize the data presented here to further define mechanisms of bacterial persistence. Given the apparent lack of overlap between mutants hypersusceptible to particular drugs we suggest that individual drugs may have individual pathways by which bacteria persist. Therefore

work should focus on understanding persistence one drug at a time. Understanding bacterial persistence may help in the design of novel compounds to accelerate therapy for mycobacterial infections - which could profoundly accelerate control efforts globally and reduce logistical costs. Innovations in TB control are critical as millions of lives continue to be lost yearly to Mtb.

Appendix A

Gene predictions for Mav

The table on the next page reports predictions for the impact of disrupting each of the annotated genomic features in Mav strain MAC109 based on my collected Tn-seq data. Each feature is labelled with the final 5 digits of the locus tag provided in the Genbank file (see Ch.2 for reference). This number is unique for each feature and only excludes an invariant alphanumeric string to distinguish the genomes from other in Genbank (ie it excludes "DFS55_" specific to the MAC109 genome). The second column provides the logarithm, base 10, of the RF (see chapter 2 for definition). The third column provides the predicted effect of disruption of each feature. See Ch.2 for definition of each two letter code and method for prediction. A computer-readable version of this table including additional information is provided as an associated file (AppendixA_FullTable.csv).

Feature	Log(RF)	Class
00005	-inf	ES
00010	-inf	GD
00015	-1.01	GD
00020	-1.13	NE
00025	-inf	ES
00030	-inf	ES
00035	0.77	NE
00040	-inf	GD
00045	-inf	GD
00050	0.39	GA
00055	0.4	NE
00060	0.45	NE
00065	0.66	NE
00070	0.66	GA
00075	0.28	NE
00080	-0.87	NE
00085	0.64	GA
00090	-inf	GD
00095	-inf	ES
00100	-inf	GD
00105	0.76	GA
00110	0.32	NE
00115	-0.53	NE
00120	0.1	NE
00125	-0.0	NE
00135	0.59	NE
00140	0.14	NE
00145	0.03	NE
00150	0.7	GA
00155	-inf	GD
00160	0.21	NE
00165	0.24	NE
00170	-0.22	NE
00175	0.42	GA
00180	0.29	GA
00185	0.11	NE
00190	0.2	NE
00195	0.26	NE
00200	-0.44	NE
00205	0.25	NE
00210	0.01	NE
00215	-inf	GD
00220	-0.57	NE
00225	-1.17	GD
00230	-0.35	NE
00235	0.29	NE
00240	0.68	NE
00245	0.15	NE
00250	0.77	GA
00255	0.07	NE
00260	0.38	NE
00265	0.72	GA
00270	0.91	NE
00275	0.5	NE
00280	-inf	ES
00285	-0.53	NE
00295	0.24	NE
00300	0.57	GA
00305	0.57	NE
00310	0.17	NE
00315	0.29	NE
00320	0.27	NE
00325	-0.57	GD
00330	-0.94	GD
00335	-0.89	GD
00340	-inf	ES
00345	0.01	NE
00350	0.73	GA
00355	0.28	NE
00360	-0.23	GD
00365	-0.42	GD
00370	0.71	NE
00375	-1.59	GD
00380	-inf	GD
00385	-inf	NE
00390	-0.08	NE
00395	-inf	ES
00400	-0.01	NE
00405	-0.25	NE
00410	-0.23	NE
00415	-0.5	GD
00420	0.45	GA
00425	-0.05	NE
00430	-0.12	NE
00435	-0.19	NE
00440	-inf	ES
00445	-0.49	NE
00450	-0.59	GD
00455	0.45	NE
00460	-0.02	NE
00465	-0.53	GD
00470	0.14	NE
00475	0.6	NE
00480	-0.7	GD
00485	-0.26	GD
00490	-1.12	GD
00495	-0.99	GD
00505	0.3	NE
00510	-0.14	NE
00515	0.09	NE
00520	0.03	NE
00525	0.19	NE
00530	0.53	GA
00535	-0.28	NE
00540	-0.08	NE
00545	0.37	NE
00550	0.11	NE
00555	-1.01	GD
00560	0.49	GA
00565	-0.26	NE
00570	-0.49	NE
00575	0.44	NE
00580	-0.95	GD
00585	-1.15	GD
00590	-1.32	GD
00595	0.0	NE
00600	-0.3	NE
00605	-0.89	GD
00610	-0.85	NE
00615	-0.38	NE
00620	-0.74	NE
00625	0.22	NE
00630	0.22	NE
00635	-1.03	NE
00640	0.19	NE
00645	-0.7	NE
00650	0.25	NE
00655	0.16	NE
00660	-0.05	NE
00665	0.74	NE
00670	-0.02	NE
00675	0.86	GA
00680	0.27	NE
00685	-0.13	NE
00690	0.36	NE
00695	0.22	NE
00700	0.13	NE
00705	-0.28	GD
00710	-0.12	NE
00715	-0.15	NE
00720	-0.23	NE
00725	0.07	NE
00730	-0.59	GD
00735	-0.18	NE
00740	0.08	NE
00745	-0.94	GD
00750	-inf	NE
00755	-0.44	NE
00760	0.73	NE
00765	-0.09	NE
00770	0.51	NE
00775	0.49	NE
00780	0.38	NE
00785	0.36	NE
00790	0.21	NE
00795	-0.44	NE
00800	0.66	NE
00805	0.26	NE
00810	-0.05	NE
00815	-0.01	NE
00820	-inf	NE
00825	0.42	NE
00830	0.04	NE
00835	-2.23	NE
00840	-0.98	GD
00845	-0.63	NE
00850	-1.95	GD
00855	-0.48	NE
00860	0.07	NE
00865	0.0	NE
00870	-0.52	GD
00875	-0.8	GD
00880	-0.35	GD
00885	-0.87	GD
00890	-0.33	GD
00895	-0.25	NE
00900	0.26	NE
00905	-0.55	GD
00910	-0.34	NE
00915	-0.57	NE
00920	0.12	NE
00925	0.01	NE
00930	-0.41	NE
00935	-0.57	NE
00940	-0.22	NE
00945	0.34	NE
00950	-0.76	GD
00955	-0.26	GD
00960	-0.36	NE
00965	-0.7	NE
00970	0.81	NE
00975	0.53	GA
00980	0.42	NE
00985	0.59	NE
00990	0.52	NE
00995	-1.29	GD
01000	-0.83	NE
01005	0.31	NE
01010	0.56	NE
01015	0.3	NE
01020	0.8	NE
01025	0.04	NE
01030	-0.23	NE
01035	-0.6	GD
01040	-0.34	NE
01045	-1.0	GD
01050	0.21	NE
01055	0.92	NE
01060	-0.65	NE
01065	0.3	NE
01070	0.03	NE
01075	-0.13	NE
01080	-0.89	NE
01085	0.3	NE
01090	-0.25	NE
01095	-0.45	NE
01100	0.39	NE
01105	-0.46	GD
01110	-0.18	NE
01115	0.53	NE
01120	0.07	NE
01125	0.06	NE
01130	0.2	NE
01135	0.34	NE
01140	0.12	NE
01145	0.3	GA
01150	0.67	GA
01155	0.59	GA
01160	-0.1	NE
01165	0.01	NE
01170	-0.13	NE
01175	-0.43	NE
01180	0.12	NE
01185	0.61	GA
01190	0.82	NE
01195	0.47	NE
01200	1.0	NE
01205	0.27	NE
01210	0.86	NE
01215	0.81	GA
01220	0.08	NE
01225	0.49	NE
01230	-0.26	NE
01235	-0.01	NE
01240	0.26	NE
01245	0.15	NE
01250	0.41	NE
01255	0.23	NE
01260	-1.08	GD
01265	-0.08	NE
01270	0.34	NE
01275	0.34	NE
01280	-1.03	NE
01285	-0.2	GD
01290	-0.88	GD
01295	-0.51	GD
01300	-1.01	GD
01305	-inf	NE
01310	-0.72	GD
01315	-0.87	GD
01320	0.08	NE
01325	0.57	NE
01330	-inf	ES
01335	0.84	NE
01340	0.51	NE
01345	-0.86	GD
01350	-1.4	GD
01355	-0.3	NE
01360	-1.62	GD
01365	-1.62	GD
01370	-1.51	GD
01375	-0.87	GD
01380	-1.03	GD
01385	-inf	GD
01390	-1.97	GD
01395	-1.33	GD
01400	-1.81	GD
01405	-2.23	GD
01410	0.06	NE
01415	-0.58	GD
01420	-0.86	NE
01425	-1.49	NE
01430	-1.16	GD
01435	-0.69	GD
01440	0.03	NE
01445	0.38	NE
01455	-0.25	NE
01460	-1.51	NE
01465	-1.01	GD
01470	-0.79	NE
01475	-0.24	GD
01480	-1.57	GD
01485	-1.0	GD
01490	-1.0	GD
01495	0.83	NE
01500	0.73	GA
01505	0.39	GA
01510	-0.66	GD
01515	-0.98	GD
01520	-0.52	GD
01525	-0.17	NE
01530	-0.25	GD
01535	-0.4	GD
01540	-0.89	NE
01545	-0.54	NE
01550	-0.19	NE
01555	-0.03	NE
01560	0.19	NE
01565	0.41	NE
01570	0.04	NE
01575	0.84	NE
01580	0.72	GA
01585	0.41	GA
01590	-inf	ES
01595	0.34	NE
01600	0.03	NE
01605	0.09	NE
01610	-0.19	NE
01615	-0.74	GD
01620	0.88	GA
01625	-0.15	NE
01630	-0.19	NE
01635	0.06	NE
01640	-0.27	NE
01645	-0.36	NE
01650	-0.98	GD
01655	-0.05	NE
01660	-0.24	NE
01665	-0.23	NE
01670	0.18	NE
01675	0.04	NE
01680	-0.11	NE
01685	-0.23	NE
01690	0.73	NE
01695	0.56	GA
01700	-inf	ES
01705	-inf	ES
01710	-inf	GD
01715	-inf	ES
01720	-inf	GD
01725	-0.01	NE
01730	0.09	NE
01735	0.61	GA
01740	0.42	GA
01745	0.23	NE
01750	-0.67	NE
01755	-0.19	NE
01760	-0.18	NE
01765	0.25	NE
01770	-0.64	NE
01775	-0.71	GD
01780	-1.02	NE
01785	-0.52	GD
01790	-0.88	GD
01795	-0.14	NE
01800	-0.9	GD
01805	-inf	ES
01810	-inf	GD
01815	-0.13	NE
01820	0.17	NE
01825	0.46	NE
01830	-0.24	NE
01835	-0.32	NE
01840	0.17	NE
01845	0.45	NE
01850	-0.07	NE
01855	-0.02	NE
01865	-0.15	NE
01870	0.4	GA
01875	0.15	NE
01880	0.1	NE
01885	0.27	NE
01890	-0.44	GD
01895	-0.43	NE
01900	-0.07	NE
01905	-0.59	GD
01910	-0.94	NE
01915	-0.18	NE
01920	-inf	NE
01925	-0.81	NE
01930	-0.47	GD
01935	-0.84	NE
01940	0.19	NE
01945	0.86	GA

Feature	Log(RF)	Class
01950	0.1	NE
01955	0.56	GA
01970	0.22	NE
01975	-0.15	NE
01980	0.14	NE
01985	-0.41	NE
01990	-0.09	NE
01995	-0.29	GD
02000	-0.42	NE
02005	-0.6	GD
02010	-0.17	NE
02015	-0.05	NE
02020	-0.08	NE
02025	-0.34	NE
02030	0.28	NE
02035	-0.27	NE
02040	0.1	NE
02045	0.03	NE
02050	0.3	NE
02055	0.02	NE
02060	-0.67	GD
02065	-0.71	GD
02070	-0.47	NE
02075	-1.28	GD
02080	-1.2	GD
02085	-0.0	NE
02090	-1.04	GD
02095	-1.29	GD
02100	-1.02	NE
02105	-1.51	GD
02110	-0.53	GD
02115	-0.18	GD
02120	-0.85	GD
02125	-0.35	NE
02130	-0.48	NE
02135	0.74	NE
02140	-1.13	GD
02145	0.28	NE
02150	-0.36	NE
02155	0.26	NE
02160	0.01	NE
02165	-0.06	NE
02170	0.3	NE
02175	0.2	NE
02180	0.74	NE
02185	0.55	NE
02190	-0.01	NE
02195	0.48	NE
02200	-0.38	GD
02205	0.16	NE
02210	-0.13	NE
02215	-1.23	GD
02220	-1.33	GD
02225	-0.74	GD
02230	0.41	NE
02235	0.37	NE
02240	-0.45	NE
02245	-0.99	GD
02250	0.09	NE
02255	-0.92	GD
02260	0.74	NE
02265	0.8	NE
02270	0.69	GA
02275	0.1	NE
02280	0.29	NE
02285	0.4	NE
02290	-0.16	NE
02295	0.43	NE
02300	0.09	NE
02305	1.04	GA
02310	0.73	NE
02315	-inf	ES
02320	0.2	NE
02325	-inf	GD
02330	-1.75	ES
02335	0.22	NE
02340	0.07	NE
02345	0.54	GA
02350	-0.6	GD
02355	0.53	GA
02360	0.67	GA
02365	0.74	GA
02370	0.26	NE
02375	-inf	ES
02380	-1.81	NE
02385	-0.28	GD
02390	0.35	NE
02395	-0.02	NE
02400	0.71	GA
02405	-0.35	NE
02410	-0.3	NE
02415	-0.93	NE
02420	-0.63	NE
02425	-0.11	NE
02430	-0.19	GD
02435	-0.3	GD
02440	-0.16	NE
02445	0.09	NE
02450	0.2	NE
02455	-inf	GD
02460	-1.46	GD
02465	-inf	GD
02470	0.12	NE
02475	0.32	NE
02480	-0.09	NE
02485	0.45	NE
02490	-inf	ES
02495	-0.09	NE
02500	0.09	NE
02505	0.52	NE
02510	0.34	NE
02515	-0.85	NE
02520	0.34	NE
02525	-inf	ES
02530	-0.18	NE
02535	0.84	NE
02540	-0.51	NE
02545	-0.06	NE
02550	0.14	NE
02555	0.05	NE
02560	-0.13	NE
02565	-0.33	NE
02570	-0.23	GD
02575	-1.0	GD
02580	-1.23	NE
02585	0.92	NE
02590	-0.29	NE
02595	-0.18	NE
02600	-0.35	GD
02605	0.03	NE
02610	-0.8	GD
02615	-0.57	NE
02620	-1.03	GD
02625	-0.34	GD
02630	-0.6	GD
02635	-0.12	NE
02640	0.13	NE
02645	0.47	NE
02650	-0.06	NE
02655	0.03	NE
02660	0.25	GA
02665	0.34	NE
02670	-2.11	GD
02675	-inf	ES
02680	-inf	GD
02685	-inf	GD
02690	-0.88	GD
02695	-0.36	NE
02700	-0.35	NE
02705	-0.53	NE
02710	0.44	NE
02715	-1.19	GD
02720	-1.41	GD
02725	0.01	NE
02730	0.12	NE
02735	-0.99	GD
02740	-0.94	GD
02745	-1.49	NE
02750	-1.71	GD
02755	-inf	GD
02760	-inf	ES
02765	-0.07	NE
02770	-0.02	NE
02780	-0.3	GD
02785	-0.96	GD
02790	-inf	GD
02795	-inf	GD
02800	0.93	GA
02805	0.07	NE
02810	-0.75	GD
02815	-inf	GD
02820	-0.66	GD
02825	-1.19	NE
02830	-1.21	GD
02835	-0.12	NE
02840	0.2	NE
02845	0.42	NE
02850	0.72	GA
02855	0.32	NE
02860	0.43	NE
02865	0.08	NE
02870	0.53	GA
02875	0.6	GA
02880	-inf	ES
02885	0.75	GA
02890	0.27	NE
02895	0.36	NE
02900	0.43	NE
02905	0.92	GA
02910	1.17	GA
02915	-0.06	NE
02920	0.36	NE
02925	-1.18	GD
02930	-0.24	NE
02935	0.22	NE
02940	0.62	NE
02945	0.58	GA
02950	0.36	NE
02955	0.35	NE
02960	0.2	NE
02965	-0.48	GD
02970	0.1	NE
02975	-0.28	NE
02985	0.52	GA
02990	-1.38	GD
02995	0.99	NE
03000	-0.37	NE
03005	-0.23	NE
03010	0.55	NE
03015	-0.45	NE
03020	0.65	NE
03025	0.43	NE
03030	0.17	NE
03035	0.18	NE
03040	0.31	NE
03045	0.64	GA
03050	1.32	GA
03055	0.43	GA
03060	0.96	GA
03065	0.29	GA
03070	0.95	NE
03075	0.53	NE
03080	0.12	NE
03085	0.53	GA
03090	0.37	NE
03095	-inf	GD
03100	0.38	GA
03105	-0.68	NE
03110	0.24	NE
03115	0.19	NE
03120	-0.61	NE
03125	0.08	NE
03130	-inf	ES
03135	-0.07	NE
03140	-1.01	GD
03145	0.96	NE
03150	-0.07	NE
03155	-0.41	NE
03160	0.19	NE
03165	-0.7	NE
03170	0.5	NE
03175	0.16	NE
03180	-1.42	GD
03185	0.29	NE
03190	0.83	GA
03195	0.36	GA
03200	0.53	GA
03205	0.8	GA
03210	-1.93	GD
03215	0.44	GA
03220	0.37	NE
03225	-inf	GD
03230	-0.34	NE
03235	-0.36	NE
03240	0.97	GA
03245	1.02	NE
03250	-inf	ES
03255	-inf	GD
03260	-inf	GD
03265	-inf	ES
03270	0.0	NE
03275	1.35	NE
03280	0.62	NE
03285	0.8	GA
03290	-0.26	NE
03295	-0.28	NE
03300	0.11	NE
03310	-0.04	NE
03315	-0.53	GD
03320	0.59	NE
03325	-1.32	GD
03330	-0.39	NE
03335	0.52	NE
03340	0.81	NE
03345	-inf	ES
03350	-1.95	GD
03355	-1.81	GD
03360	-2.27	GD
03365	-1.79	GD
03370	-inf	GD
03375	0.49	GA
03380	0.55	NE
03385	-1.27	NE
03390	-0.12	NE
03395	-1.2	GD
03400	-0.33	NE
03405	-1.19	GD
03410	-0.87	GD
03415	0.31	NE
03420	-2.11	GD
03425	0.09	NE
03430	-1.28	GD
03435	-0.89	NE
03440	-0.48	NE
03445	0.42	NE
03450	0.33	NE
03455	-0.63	NE
03460	0.85	GA
03465	-1.07	GD
03470	-0.62	GD
03475	0.35	NE
03480	0.25	NE
03485	0.87	GA
03490	-0.06	NE
03495	-0.86	NE
03500	-0.21	NE
03505	-1.29	GD
03510	-inf	GD
03515	0.53	NE
03520	-1.93	NE
03525	-1.08	GD
03530	0.01	NE
03535	0.19	NE
03540	-1.01	GD
03545	0.38	GA
03550	-0.35	NE
03555	0.4	NE
03560	0.49	NE
03565	-0.15	NE
03570	0.87	GA
03575	-inf	GD
03580	0.54	NE
03585	0.44	NE
03590	-0.38	NE
03595	0.15	NE
03600	-0.04	NE
03605	0.01	NE
03610	0.55	GA
03615	0.37	NE
03620	-0.17	NE
03625	-0.18	NE
03630	0.4	NE
03635	0.18	NE
03640	0.14	NE
03645	0.21	NE
03650	-0.75	GD
03655	0.15	NE
03660	0.31	NE
03665	0.14	NE
03670	-0.1	NE
03675	-0.95	GD
03680	0.37	NE
03685	-0.82	GD
03690	-0.98	NE
03695	-0.71	NE
03700	-0.72	NE
03705	-0.96	GD
03710	-0.94	GD
03715	-0.51	GD
03720	-1.62	GD
03725	-1.93	ES
03730	-0.93	GD
03735	-1.03	GD
03740	-0.74	GD
03745	-0.78	GD
03750	-1.2	GD
03755	-0.79	GD
03760	-1.02	GD
03765	0.59	GA
03770	-0.08	NE
03775	-0.31	NE
03780	-0.24	NE
03785	-0.15	NE
03790	0.06	NE
03795	-0.56	NE
03800	0.35	NE
03805	0.68	GA
03810	0.07	NE
03815	0.19	NE
03820	-inf	ES
03825	-inf	ES
03830	-0.58	NE
03835	-inf	ES
03840	-inf	ES
03845	-inf	NE
03850	-inf	NE
03855	-inf	GD
03860	-inf	ES
03865	-inf	GD
03870	-inf	GD
03875	0.48	NE
03880	0.12	NE
03885	0.51	GA
03890	0.27	GA

Feature	Log(RF)	Class
03895	0.62	GA
03900	-0.12	NE
03905	0.73	NE
03910	-0.02	NE
03915	-0.73	NE
03920	-1.11	NE
03925	0.05	NE
03930	-inf	GD
03935	0.58	NE
03940	0.69	NE
03945	0.06	NE
03950	-1.64	GD
03955	-inf	ES
03960	-inf	ES
03965	-inf	ES
03970	-inf	ES
03975	0.57	GA
03980	0.9	GA
03985	0.79	NE
03990	-0.14	NE
03995	-1.26	GD
04000	-1.1	GD
04005	-1.17	GD
04010	-0.42	NE
04015	-0.64	NE
04020	0.12	NE
04025	-inf	GD
04030	-inf	GD
04035	-inf	ES
04040	-inf	ES
04045	0.1	NE
04050	0.61	NE
04055	-0.55	NE
04060	-0.32	NE
04065	-0.29	NE
04070	-0.15	NE
04075	-1.13	NE
04080	-0.96	NE
04085	0.01	NE
04090	0.16	NE
04095	-0.86	GD
04100	-0.62	GD
04105	-0.76	GD
04110	-0.52	GD
04115	-1.16	GD
04120	0.35	NE
04125	-inf	ES
04130	-inf	ES
04135	-inf	ES
04140	-inf	GD
04145	-inf	GD
04155	-inf	GD
04160	-inf	GD
04165	-inf	ES
04170	-inf	GD
04175	-1.34	NE
04180	-0.0	NE
04185	-0.27	GD
04190	-1.02	GD
04195	-0.21	GD
04200	0.45	NE
04205	-0.41	NE
04210	-inf	GD
04215	-inf	NE
04220	-inf	GD
04225	-inf	NE
04230	-inf	ES
04235	-inf	ES
04240	-inf	GD
04245	-inf	GD
04250	-inf	GD
04255	-1.58	GD
04260	-0.16	NE
04265	-0.68	GD
04270	0.02	NE
04275	-0.21	NE
04280	-0.49	NE
04285	-0.21	NE
04290	-0.09	NE
04295	-0.0	NE
04300	-0.04	NE
04305	-1.19	GD
04310	-0.67	NE
04315	0.12	NE
04320	-inf	ES
04325	-inf	GD
04330	-0.76	NE
04335	1.0	GA
04340	-0.24	NE
04345	-0.14	NE
04350	-1.12	GD
04355	-0.14	NE
04360	0.3	NE
04365	-0.22	NE
04370	-1.75	GD
04375	-0.21	GD
04380	-0.57	GD
04385	-0.32	NE
04390	-0.6	NE
04395	-0.28	NE
04400	-0.52	GD
04405	-0.09	NE
04410	0.09	NE
04415	-0.61	NE
04420	-0.69	NE
04425	0.67	NE
04430	-0.33	NE
04435	-1.49	NE
04440	0.3	NE
04445	-0.13	NE
04450	0.45	GA
04455	-inf	GD
04460	-inf	ES
04465	0.1	NE
04470	-0.1	NE
04475	-inf	ES
04485	-inf	ES
04490	-1.97	NE
04495	-inf	ES
04500	-inf	GD
04505	-inf	GD
04510	-1.93	GD
04515	-1.02	NE
04520	0.32	NE
04525	0.37	NE
04530	0.41	NE
04535	-1.45	GD
04540	-1.06	GD
04545	-0.45	NE
04550	0.66	NE
04555	-0.37	NE
04560	-inf	ES
04565	-1.67	GD
04570	-inf	ES
04575	-0.1	NE
04580	-0.53	NE
04585	0.6	GA
04590	0.04	NE
04595	0.78	NE
04600	-0.07	NE
04605	-inf	ES
04610	0.64	NE
04615	0.06	NE
04620	0.16	NE
04625	0.38	GA
04630	-inf	ES
04635	-0.62	GD
04640	-1.64	GD
04645	-0.8	GD
04650	-0.44	GD
04655	-1.11	GD
04660	-inf	GD
04665	0.33	NE
04670	0.13	NE
04675	-0.36	NE
04680	0.26	NE
04685	0.38	NE
04690	0.53	NE
04695	-0.08	NE
04700	0.33	NE
04705	-inf	ES
04710	-1.46	GD
04715	0.36	NE
04720	0.67	GA
04725	-inf	NE
04730	0.37	GA
04735	0.55	NE
04740	-0.47	GD
04745	-inf	ES
04750	0.56	NE
04755	-0.5	GD
04760	-0.41	NE
04765	-0.27	NE
04770	-0.0	NE
04775	-0.37	GD
04780	0.23	NE
04785	0.17	NE
04790	0.27	NE
04795	-inf	GD
04800	-1.86	GD
04805	-0.02	NE
04810	0.56	NE
04815	0.25	NE
04820	0.36	NE
04825	0.5	GA
04830	0.29	NE
04835	0.61	GA
04840	-0.39	NE
04845	-0.35	NE
04850	0.73	GA
04855	0.4	NE
04860	0.57	GA
04865	1.0	NE
04870	-inf	GD
04875	-inf	NE
04880	1.19	NE
04885	-0.41	NE
04890	-0.96	GD
04895	-1.24	GD
04900	-0.8	GD
04905	-1.2	GD
04910	0.14	NE
04915	0.1	NE
04920	-inf	ES
04925	-0.09	NE
04930	-0.31	NE
04935	-1.64	GD
04940	-0.87	GD
04945	-0.34	NE
04950	0.18	NE
04955	-0.28	NE
04960	-1.09	GD
04965	0.21	NE
04970	0.3	NE
04975	-0.19	NE
04980	0.54	NE
04985	0.09	NE
04990	-0.03	NE
04995	0.45	NE
05000	-0.2	NE
05005	0.25	NE
05010	0.65	NE
05015	0.37	NE
05020	0.23	NE
05025	0.39	NE
05030	0.13	NE
05035	-0.71	GD
05040	-0.28	GD
05045	-0.22	NE
05050	0.12	NE
05055	0.38	GA
05060	0.07	NE
05065	-0.01	NE
05070	0.07	NE
05075	-1.62	GD
05080	-0.74	NE
05085	-0.14	NE
05090	0.47	NE
05095	0.42	NE
05100	-0.34	GD
05105	-0.5	NE
05110	0.13	NE
05115	0.29	NE
05120	-0.22	NE
05125	-0.29	GD
05130	0.68	GA
05135	0.17	NE
05140	-0.31	GD
05145	-0.01	NE
05150	0.12	NE
05155	0.97	NE
05160	0.51	NE
05165	0.85	NE
05170	1.0	GA
05175	0.51	NE
05180	-inf	ES
05185	0.69	NE
05190	0.59	GA
05195	0.06	NE
05200	0.86	NE
05205	-inf	ES
05210	0.46	GA
05215	-inf	NE
05220	0.81	NE
05225	-inf	GD
05230	-inf	GD
05235	-2.11	GD
05240	0.65	NE
05245	0.98	GA
05250	0.31	NE
05255	0.38	NE
05260	0.57	GA
05265	0.35	GA
05270	-0.05	NE
05275	0.47	NE
05280	-0.97	GD
05285	0.45	NE
05290	0.25	NE
05295	0.0	NE
05300	-2.23	ES
05305	-inf	ES
05310	-inf	ES
05315	1.0	NE
05320	0.09	NE
05325	0.09	NE
05330	-1.36	NE
05335	-0.36	NE
05340	-inf	NE
05345	0.43	NE
05350	-1.18	NE
05355	-inf	ES
05360	0.08	NE
05365	-inf	ES
05370	0.23	NE
05375	0.58	NE
05380	0.49	GA
05385	-0.25	NE
05390	0.49	NE
05395	0.99	GA
05400	-inf	ES
05405	-inf	ES
05410	-1.41	GD
05415	-0.59	NE
05420	-1.37	GD
05425	-0.51	NE
05430	0.66	GA
05435	-inf	ES
05440	0.47	NE
05445	0.11	NE
05450	-0.47	GD
05455	0.01	NE
05460	0.24	NE
05465	-0.1	NE
05470	0.79	GA
05475	0.5	GA
05480	0.8	GA
05485	-0.0	NE
05490	-0.55	NE
05495	-0.18	NE
05500	-0.14	NE
05505	0.28	NE
05510	-0.28	GD
05515	-0.7	GD
05520	-0.88	GD
05525	-0.38	GD
05530	0.15	NE
05535	0.17	NE
05540	-inf	GD
05545	0.33	NE
05550	0.21	GA
05555	1.27	NE
05560	0.2	NE
05565	-0.18	NE
05570	-0.95	GD
05575	-1.24	GD
05580	0.07	NE
05585	-0.3	NE
05590	-0.82	NE
05595	0.82	GA
05600	0.46	NE
05610	0.5	NE
05615	-0.01	NE
05620	-1.1	GD
05625	0.66	NE
05630	0.44	NE
05635	-0.21	NE
05640	-1.29	GD
05645	0.12	NE
05650	0.15	NE
05655	0.35	GA
05660	0.3	GA
05665	0.79	GA
05670	0.31	NE
05675	0.97	NE
05680	-0.86	GD
05685	0.08	NE
05690	-inf	ES
05695	-0.3	NE
05700	-0.9	NE
05705	-0.55	GD
05710	0.1	NE
05715	-0.4	NE
05720	-0.96	GD
05725	0.0	NE
05730	0.09	NE
05735	0.83	NE
05740	-inf	ES
05745	0.78	NE
05750	1.04	NE
05755	0.36	NE
05760	0.58	NE
05765	-0.28	NE
05770	-0.23	NE
05775	-0.42	GD
05780	-inf	GD
05785	0.32	NE
05790	-0.87	GD
05795	-0.56	GD
05800	0.19	NE
05805	0.16	NE
05810	0.04	NE
05815	0.29	NE
05820	-0.24	NE
05825	0.3	NE

Feature	Log(RF)	Class
05830	0.12	NE
05835	-0.39	GD
05840	-0.06	NE
05845	0.69	GA
05850	0.92	GA
05855	0.62	GA
05860	-0.07	NE
05865	-0.14	NE
05870	0.19	NE
05875	0.19	NE
05880	-0.21	NE
05885	0.06	NE
05890	0.07	NE
05895	0.43	NE
05900	-0.14	NE
05905	0.39	GA
05915	0.4	GA
05920	-0.04	NE
05925	-0.29	NE
05930	0.69	NE
05935	0.05	NE
05940	0.4	GA
05945	0.63	GA
05950	0.45	NE
05955	0.08	NE
05960	0.37	GA
05965	0.24	NE
05970	0.15	NE
05975	-0.04	NE
05980	0.18	NE
05985	0.28	NE
05990	0.63	GA
05995	-0.06	NE
06000	-0.17	NE
06005	0.77	NE
06010	0.04	NE
06015	0.42	GA
06020	0.49	GA
06025	0.79	GA
06030	0.61	GA
06035	1.02	GA
06040	0.34	GA
06045	0.6	NE
06050	-0.41	NE
06055	-0.49	GD
06060	-1.56	GD
06065	-0.51	GD
06070	-0.26	GD
06075	-0.33	NE
06080	-0.69	GD
06085	-0.5	NE
06090	-0.37	NE
06095	0.01	NE
06100	0.69	NE
06105	0.34	NE
06110	-0.11	NE
06115	0.52	NE
06120	-0.45	NE
06125	0.03	NE
06130	-0.2	NE
06135	0.69	GA
06140	0.27	NE
06145	0.27	NE
06150	0.04	NE
06155	0.9	GA
06160	0.53	GA
06165	0.28	NE
06170	0.22	GA
06175	-0.08	NE
06185	-1.5	GD
06190	-0.41	GD
06195	-0.38	GD
06200	-inf	NE
06205	-0.23	NE
06210	-0.47	GD
06215	-1.27	GD
06220	-0.86	NE
06225	-inf	GD
06230	-0.92	GD
06235	-inf	ES
06240	-0.39	NE
06245	-0.23	NE
06250	-0.77	GD
06255	-0.6	GD
06260	-0.02	NE
06265	-0.05	NE
06270	0.3	NE
06275	-inf	NE
06280	-inf	NE
06285	-inf	GD
06290	-1.93	GD
06295	-0.65	GD
06300	-1.08	NE
06305	-0.21	GD
06310	0.48	GA
06315	0.51	NE
06320	0.32	NE
06325	0.66	NE
06330	-0.4	NE
06335	-0.41	GD
06340	-0.97	NE
06345	-0.38	GD
06350	-0.36	GD
06355	-0.06	NE
06360	-0.47	NE
06365	-0.73	GD
06370	-0.92	GD
06375	-0.93	GD
06380	-0.11	NE
06385	-0.32	NE
06390	-0.85	NE
06395	0.07	NE
06400	-0.78	GD
06405	-0.87	GD
06410	0.21	NE
06415	-0.93	GD
06420	-1.31	GD
06425	-1.21	GD
06430	-1.62	GD
06435	-1.2	GD
06440	-0.63	GD
06445	-1.45	GD
06450	-0.88	GD
06455	-0.93	GD
06460	-1.33	GD
06465	-1.22	GD
06470	-0.22	GD
06475	-1.81	ES
06480	-0.93	GD
06485	-0.43	GD
06490	-0.41	NE
06495	-0.66	NE
06500	-0.31	NE
06505	-1.07	NE
06510	-1.0	NE
06515	0.12	NE
06520	-1.54	GD
06525	-0.58	GD
06530	-1.04	GD
06535	-1.17	GD
06540	-0.77	GD
06545	-0.63	GD
06550	-0.56	NE
06555	-0.73	GD
06560	-0.4	NE
06565	-inf	GD
06570	0.31	NE
06575	-0.88	GD
06580	-0.28	GD
06585	-1.1	GD
06590	-1.42	GD
06595	0.08	NE
06600	-1.34	GD
06605	-0.18	NE
06610	0.55	NE
06615	0.7	GA
06620	-inf	ES
06625	-0.11	NE
06630	0.84	GA
06635	-inf	ES
06640	-1.93	ES
06645	-inf	ES
06650	0.1	NE
06655	0.95	GA
06660	-inf	ES
06665	0.28	NE
06670	0.63	NE
06675	-0.37	GD
06680	0.29	NE
06685	0.31	NE
06690	0.05	NE
06695	-0.74	GD
06700	-0.17	NE
06705	0.4	NE
06710	-0.55	GD
06715	0.22	NE
06720	-0.17	NE
06725	-0.87	NE
06730	-0.06	NE
06735	-0.25	NE
06740	0.54	NE
06745	-0.95	GD
06750	0.27	NE
06755	-0.6	GD
06760	0.94	NE
06765	0.09	NE
06770	0.47	GA
06775	0.16	NE
06780	-inf	ES
06785	-0.61	GD
06790	0.58	NE
06795	0.65	NE
06800	0.67	NE
06805	0.25	NE
06810	-0.28	GD
06815	-0.59	NE
06820	-0.28	NE
06825	-0.16	NE
06830	-0.56	GD
06835	-0.71	GD
06840	-0.66	GD
06845	-1.81	GD
06850	-0.49	GD
06855	-1.23	GD
06860	-0.1	NE
06865	-inf	GD
06870	-0.58	GD
06875	0.5	NE
06880	-0.15	NE
06885	0.53	NE
06890	-1.07	GD
06895	-0.57	NE
06900	-0.48	NE
06905	-0.74	GD
06910	-0.24	NE
06915	-0.12	NE
06920	-0.75	GD
06925	-0.45	NE
06930	0.32	NE
06935	0.35	NE
06940	-0.18	NE
06945	-0.47	NE
06950	0.38	NE
06955	-0.36	NE
06960	-0.55	NE
06965	0.11	NE
06970	0.36	NE
06975	0.2	NE
06980	0.13	NE
06985	0.31	NE
06990	0.53	GA
06995	0.6	NE
07000	0.44	GA
07005	0.38	NE
07010	0.49	NE
07015	0.38	NE
07020	-0.07	NE
07025	-inf	ES
07030	-inf	ES
07035	-inf	ES
07040	0.12	NE
07045	0.08	NE
07050	0.22	NE
07055	-0.18	NE
07060	-inf	ES
07065	0.35	NE
07070	0.61	GA
07075	-1.93	GD
07080	-inf	ES
07085	-inf	GD
07090	-0.01	NE
07095	0.29	NE
07100	-0.12	NE
07105	-0.32	NE
07110	-0.26	NE
07115	0.57	GA
07120	0.93	NE
07125	-inf	ES
07130	-0.07	NE
07140	-1.15	GD
07145	0.05	NE
07150	0.49	NE
07155	-inf	GD
07160	-0.36	NE
07165	-0.3	NE
07170	-0.98	GD
07175	-0.7	GD
07180	-0.82	GD
07185	0.04	NE
07190	-0.05	NE
07195	-0.69	GD
07200	-0.05	NE
07205	0.12	NE
07210	0.25	NE
07215	-0.66	GD
07220	-1.62	GD
07225	-0.98	GD
07230	-0.33	NE
07235	-0.52	GD
07240	0.03	NE
07245	-inf	GD
07250	-1.97	GD
07255	0.0	NE
07260	0.33	NE
07265	-1.15	GD
07270	-1.24	GD
07275	0.22	NE
07280	0.11	NE
07285	-inf	ES
07290	0.17	NE
07295	-0.17	NE
07300	0.17	NE
07305	-0.7	NE
07310	-0.53	GD
07315	-0.78	GD
07320	-0.53	NE
07330	-inf	ES
07335	-0.61	GD
07340	-0.42	NE
07345	-inf	ES
07350	0.51	NE
07355	-0.03	NE
07360	-0.13	NE
07365	-0.68	GD
07370	-1.62	NE
07375	-1.24	GD
07380	-0.03	NE
07385	-0.12	NE
07390	-inf	GD
07395	-0.77	GD
07400	-0.81	GD
07405	-1.34	GD
07410	-inf	ES
07415	-inf	ES
07420	-inf	GD
07425	0.16	NE
07430	-0.39	GD
07435	-0.87	GD
07440	-1.42	GD
07445	-1.15	GD
07450	-0.3	GD
07455	-0.81	GD
07460	-0.16	NE
07465	0.62	NE
07470	-inf	ES
07475	-1.81	GD
07480	0.05	NE
07485	-0.52	NE
07490	-0.58	NE
07495	0.11	NE
07500	0.64	GA
07505	0.58	GA
07510	-0.38	NE
07520	-0.33	NE
07525	0.29	NE
07530	-0.7	NE
07535	0.33	NE
07540	-inf	GD
07550	0.08	NE
07555	-inf	ES
07560	-inf	GD
07565	0.12	NE
07570	-inf	GD
07575	-0.24	NE
07580	0.4	NE
07585	0.41	GA
07590	0.26	NE
07595	0.22	NE
07600	0.77	GA
07605	0.8	GA
07610	0.63	GA
07615	0.59	GA
07620	0.39	NE
07625	0.62	GA
07630	0.82	GA
07635	0.26	NE
07640	0.98	GA
07645	-inf	ES
07650	0.37	GA
07655	-0.8	GD
07660	0.21	NE
07665	0.35	GA
07670	0.43	NE
07675	0.43	GA
07680	0.57	GA
07685	0.37	GA
07690	0.43	GA
07695	0.01	NE
07700	0.17	NE
07705	-1.32	GD
07710	-0.04	NE
07715	-0.54	GD
07720	-0.03	NE
07725	0.15	NE
07730	-0.07	NE
07735	0.3	NE
07745	0.32	NE
07750	-inf	GD
07755	-inf	NE
07760	-inf	GD
07765	-0.21	NE
07770	-0.78	GD
07775	-0.32	NE
07780	-inf	GD

Feature	Log(RF)	Class
07785	0.0	NE
07790	-inf	GD
07795	-0.04	NE
07800	-0.14	NE
07805	-0.61	GD
07810	0.4	NE
07815	0.57	NE
07820	-inf	NE
07825	0.13	NE
07830	-0.07	NE
07835	-0.01	NE
07840	0.33	NE
07845	0.66	GA
07850	-inf	GD
07855	-0.51	NE
07860	-0.91	GD
07865	-inf	GD
07870	0.27	NE
07875	0.44	NE
07880	-inf	GD
07885	0.42	NE
07890	0.69	GA
07895	0.77	GA
07900	0.79	GA
07905	0.48	NE
07910	-0.31	NE
07915	-0.24	NE
07920	-0.27	NE
07925	-0.23	NE
07930	0.72	GA
07935	0.84	GA
07940	0.26	NE
07945	0.68	GA
07950	0.57	NE
07955	-inf	GD
07960	-inf	GD
07965	0.39	NE
07970	-0.15	NE
07975	0.83	NE
07980	0.03	NE
07985	0.69	GA
07990	0.29	NE
07995	-0.27	NE
08000	0.36	NE
08005	-inf	ES
08010	0.22	NE
08015	-inf	GD
08020	-2.23	GD
08025	0.54	GA
08030	0.63	GA
08035	0.54	GA
08040	0.14	NE
08045	-inf	ES
08050	-0.74	GD
08055	-inf	GD
08060	0.32	NE
08065	0.6	GA
08070	0.62	NE
08075	-inf	ES
08080	-0.98	GD
08085	-0.33	NE
08090	-0.66	GD
08095	-0.86	GD
08100	-1.34	GD
08105	-0.46	NE
08110	-0.41	NE
08115	-0.15	NE
08120	0.15	NE
08125	-0.39	NE
08130	-1.21	GD
08135	-0.62	GD
08140	-0.69	GD
08145	-0.64	GD
08150	-0.77	GD
08155	-0.82	NE
08160	-1.07	GD
08165	-0.69	GD
08170	-0.47	NE
08175	0.21	NE
08180	-inf	ES
08185	-0.41	NE
08190	-0.26	NE
08195	-0.45	NE
08200	-inf	GD
08205	1.08	GA
08210	-inf	ES
08215	-0.56	NE
08220	-0.66	NE
08225	0.31	NE
08230	0.25	NE
08235	0.32	NE
08240	-0.49	NE
08245	-0.45	GD
08250	0.44	GA
08255	-0.87	GD
08260	-inf	GD
08265	0.44	NE
08270	-1.24	GD
08275	-0.24	GD
08280	0.11	NE
08285	-inf	ES
08290	-0.48	NE
08295	-inf	ES
08300	-0.08	NE
08305	-0.01	NE
08310	-0.39	NE
08315	0.11	NE
08320	0.08	NE
08325	-0.64	GD
08330	-0.69	GD
08335	-0.19	NE
08340	0.43	NE
08345	0.46	NE
08350	0.09	NE
08355	-0.04	NE
08360	0.41	NE
08365	-1.28	NE
08370	0.16	NE
08375	0.17	NE
08380	-0.75	NE
08385	-0.44	GD
08390	-0.95	GD
08395	-1.34	NE
08400	0.0	NE
08405	0.15	NE
08410	-0.1	NE
08415	-inf	ES
08420	-inf	GD
08425	-0.65	NE
08430	-0.18	GD
08440	-0.48	NE
08445	-inf	GD
08450	-0.94	GD
08455	-0.62	GD
08460	-0.2	GD
08465	0.81	NE
08470	-inf	ES
08475	0.71	GA
08480	-inf	GD
08485	0.92	NE
08490	0.96	GA
08495	0.4	NE
08500	-0.12	NE
08505	0.54	NE
08510	-0.0	NE
08515	1.03	GA
08520	1.05	GA
08525	0.67	GA
08530	0.39	NE
08535	-0.37	NE
08540	0.25	NE
08545	1.23	GA
08550	0.42	NE
08555	0.99	GA
08560	-0.72	GD
08565	-0.7	GD
08570	-inf	GD
08575	0.53	GA
08580	0.43	GA
08585	0.09	NE
08590	0.29	GA
08595	-inf	NE
08600	-0.45	NE
08605	0.52	NE
08610	0.22	NE
08615	0.34	NE
08620	-0.19	NE
08625	0.7	GA
08630	0.09	NE
08635	0.76	GA
08640	0.14	NE
08645	-0.67	NE
08650	0.68	NE
08655	0.62	NE
08660	0.67	GA
08665	0.27	NE
08670	0.44	NE
08675	-0.12	NE
08680	-0.65	NE
08685	0.2	NE
08690	-inf	ES
08695	-0.25	NE
08700	0.05	NE
08705	0.25	NE
08710	-1.54	GD
08715	-1.15	GD
08720	-1.05	NE
08725	-0.76	NE
08730	-0.02	NE
08735	-0.11	NE
08740	0.55	GA
08745	0.15	NE
08750	-1.2	GD
08755	-1.62	GD
08760	-1.19	GD
08765	-inf	ES
08770	0.44	NE
08775	-inf	ES
08780	0.06	NE
08785	0.34	NE
08790	0.6	GA
08795	-inf	ES
08800	-inf	ES
08805	-inf	ES
08810	-0.58	NE
08815	-0.3	NE
08820	-0.54	NE
08825	-inf	ES
08830	0.34	NE
08835	0.42	NE
08840	-0.38	NE
08845	-1.27	GD
08850	0.22	NE
08855	0.87	NE
08860	-0.32	NE
08865	0.7	GA
08870	-0.4	NE
08875	-0.48	NE
08880	-0.04	NE
08885	-0.53	GD
08890	-0.06	NE
08895	-0.66	GD
08900	-0.7	NE
08905	-0.36	NE
08910	0.07	NE
08920	0.25	NE
08925	-0.54	NE
08930	-0.54	NE
08935	-0.17	NE
08940	-1.23	GD
08945	-1.05	GD
08950	-0.41	NE
08955	0.42	NE
08960	-0.2	NE
08965	-0.13	NE
08970	0.48	NE
08975	0.2	NE
08980	-inf	GD
08990	-inf	ES
09000	-0.67	GD
09005	0.72	NE
09010	-0.33	GD
09015	-0.46	GD
09020	-0.26	NE
09025	-0.6	NE
09030	-0.49	GD
09035	-0.45	GD
09040	0.16	NE
09045	0.2	NE
09050	0.15	NE
09055	0.4	NE
09060	-0.32	NE
09065	-0.05	NE
09070	-0.37	NE
09080	-0.86	GD
09085	0.11	NE
09090	-0.1	NE
09095	0.43	NE
09100	-1.24	GD
09105	-1.11	GD
09110	-1.38	GD
09115	0.63	NE
09120	0.42	NE
09125	-0.34	NE
09130	0.81	GA
09135	-1.62	NE
09140	0.34	NE
09145	0.78	NE
09150	0.25	NE
09155	0.08	NE
09160	0.49	GA
09165	0.1	NE
09170	-0.08	NE
09175	-0.32	NE
09180	-inf	ES
09185	-0.85	GD
09190	-inf	GD
09195	-inf	ES
09200	-inf	GD
09205	-1.93	GD
09210	0.42	NE
09215	0.17	NE
09220	-1.76	GD
09225	-1.56	GD
09230	0.01	NE
09235	0.81	GA
09240	-0.03	NE
09245	0.22	NE
09250	0.92	GA
09255	0.19	NE
09260	-0.69	NE
09265	-0.8	GD
09270	0.26	NE
09275	0.47	GA
09280	0.47	GA
09285	1.22	GA
09290	-0.89	GD
09295	-0.54	GD
09300	-0.23	NE
09305	0.84	NE
09310	-0.76	GD
09315	-1.04	GD
09320	-inf	ES
09325	-0.18	NE
09330	-inf	ES
09335	-0.37	NE
09340	-0.12	NE
09345	-0.0	NE
09350	-0.13	NE
09355	-0.77	GD
09360	0.45	NE
09365	-0.18	NE
09370	-0.15	NE
09375	-1.19	GD
09380	-0.86	GD
09385	-0.48	NE
09390	-0.56	NE
09395	-inf	ES
09400	0.34	NE
09405	-0.16	NE
09410	0.79	GA
09415	0.8	GA
09420	0.55	NE
09425	0.72	NE
09430	0.17	NE
09435	0.24	NE
09440	-0.49	NE
09445	-0.27	GD
09450	-0.07	NE
09455	0.17	NE
09460	1.04	NE
09465	0.72	NE
09470	0.51	NE
09475	-inf	ES
09485	-0.3	NE
09490	-inf	GD
09495	0.31	NE
09500	0.18	NE
09505	-0.3	NE
09510	-1.33	NE
09515	-0.07	NE
09520	-0.34	GD
09525	-0.05	NE
09530	0.4	NE
09535	0.33	NE
09540	-inf	ES
09545	-inf	ES
09550	-inf	ES
09555	-inf	GD
09560	0.27	NE
09565	-0.11	NE
09570	-inf	GD
09575	-inf	ES
09580	0.39	NE
09585	-0.3	GD
09595	-0.52	NE
09600	-0.93	GD
09605	-1.08	NE
09610	-0.72	NE
09615	-0.85	GD
09620	-1.34	GD
09625	-0.79	GD
09630	-0.42	NE
09635	-0.73	NE
09640	-0.69	GD
09645	-1.06	NE
09650	0.08	NE
09655	-1.93	GD
09660	-inf	GD
09665	-inf	GD
09670	-1.5	NE
09675	-inf	ES
09680	-inf	ES
09685	-inf	GD
09695	-0.73	NE
09700	0.48	NE
09705	-inf	GD
09710	-1.97	GD
09715	-inf	ES
09720	0.15	NE
09725	-0.22	NE
09730	-0.03	NE
09735	-0.21	NE
09740	-0.46	NE

Feature	Log(RF)	Class
09745	-0.34	NE
09750	-inf	ES
09755	-1.08	NE
09760	0.35	NE
09765	-inf	ES
09770	0.21	NE
09775	-inf	GD
09780	-inf	GD
09785	-1.45	GD
09790	-1.06	NE
09795	-inf	ES
09800	-inf	ES
09805	-inf	NE
09810	-1.23	GD
09815	0.25	NE
09820	0.38	NE
09825	-0.37	GD
09830	0.61	GA
09835	-0.34	GD
09840	-inf	GD
09845	0.35	NE
09850	-0.14	NE
09855	0.07	NE
09860	-0.64	NE
09865	-0.24	NE
09870	-0.13	NE
09875	-0.19	NE
09880	-0.83	GD
09885	0.08	NE
09890	0.24	NE
09895	-0.55	NE
09900	-0.45	NE
09905	0.22	NE
09910	-1.1	GD
09915	-1.67	GD
09920	-0.95	GD
09925	0.16	NE
09930	0.08	NE
09935	-0.76	GD
09940	-0.4	NE
09945	-0.56	GD
09950	-0.33	GD
09955	0.03	NE
09960	-1.46	NE
09965	-0.04	NE
09970	-0.59	GD
09975	-1.03	GD
09980	-inf	ES
09985	-inf	GD
09990	0.22	NE
09995	0.77	NE
10000	-1.45	GD
10005	-0.71	GD
10010	-0.07	NE
10015	-inf	GD
10020	-0.0	NE
10025	-inf	ES
10030	-inf	ES
10035	-inf	ES
10040	-1.81	GD
10045	-0.66	GD
10050	-0.75	NE
10055	0.58	NE
10060	0.64	NE
10070	-0.15	NE
10075	0.86	NE
10080	0.46	GA
10085	-0.1	NE
10090	-0.7	NE
10095	0.06	NE
10100	-inf	ES
10105	-inf	GD
10110	-1.64	GD
10115	0.45	NE
10120	0.47	NE
10125	0.28	NE
10130	0.06	NE
10135	0.02	NE
10140	-inf	GD
10145	-inf	ES
10150	-inf	ES
10155	-0.47	NE
10160	-0.55	NE
10165	0.29	NE
10170	0.61	NE
10175	0.11	NE
10185	0.69	NE
10190	-0.73	GD
10200	-0.42	NE
10205	0.88	NE
10210	-0.43	NE
10215	-0.29	NE
10220	0.1	NE
10225	0.63	GA
10230	0.96	GA
10235	0.19	NE
10240	0.31	NE
10245	0.22	NE
10250	0.46	GA
10255	0.55	GA
10260	0.44	NE
10265	-0.12	NE
10270	-0.12	NE
10275	-0.13	NE
10280	-0.42	NE
10285	-1.97	GD
10290	0.27	NE
10295	-0.52	GD
10300	0.25	NE
10305	0.21	NE
10310	0.12	NE
10315	-0.37	GD
10320	0.15	NE
10325	0.22	NE
10330	-0.48	GD
10335	-0.04	NE
10340	0.4	GA
10345	0.37	GA
10350	0.39	GA
10355	0.49	NE
10360	-0.12	NE
10365	0.29	NE
10370	0.04	NE
10375	0.22	NE
10380	-0.28	GD
10385	-0.23	NE
10390	-0.07	NE
10395	0.69	NE
10400	0.37	NE
10405	0.74	GA
10410	0.2	NE
10415	0.42	GA
10420	-0.65	GD
10425	-1.16	NE
10430	0.12	NE
10435	-0.66	NE
10440	0.22	NE
10445	-inf	ES
10450	0.05	NE
10455	0.04	NE
10460	0.5	GA
10465	-1.94	GD
10470	-0.71	NE
10475	0.64	NE
10480	0.58	NE
10485	0.43	GA
10490	0.34	NE
10495	0.07	NE
10500	-inf	ES
10505	-0.49	NE
10510	0.42	NE
10515	-0.28	NE
10520	0.03	NE
10525	0.77	GA
10530	0.81	GA
10535	0.15	NE
10540	0.31	NE
10545	-0.09	NE
10550	-1.23	GD
10555	-0.68	GD
10560	-1.43	GD
10565	-0.38	GD
10570	0.28	NE
10575	0.83	GA
10580	1.07	GA
10585	0.78	GA
10590	0.45	NE
10595	-inf	NE
10600	-1.93	GD
10605	-0.38	NE
10610	-0.54	GD
10615	0.46	GA
10620	0.8	NE
10625	0.32	NE
10630	1.13	GA
10635	0.96	GA
10640	0.89	GA
10645	0.84	NE
10650	0.45	GA
10655	-0.54	NE
10660	0.27	NE
10665	-0.26	NE
10670	0.28	NE
10675	0.45	NE
10680	-inf	GD
10685	-inf	GD
10690	-inf	GD
10695	-inf	GD
10700	-inf	GD
10705	-0.54	NE
10710	-inf	GD
10715	-inf	GD
10720	0.39	NE
10725	-1.2	NE
10730	-inf	ES
10735	-0.78	GD
10740	-inf	GD
10745	-inf	ES
10750	-inf	ES
10755	-inf	ES
10760	0.43	NE
10765	0.5	NE
10770	0.63	NE
10775	-0.99	NE
10780	-0.31	NE
10785	-0.28	NE
10790	-0.2	NE
10795	-0.6	NE
10800	0.29	NE
10805	0.56	GA
10810	-2.11	GD
10815	0.74	NE
10820	-0.24	NE
10825	0.82	GA
10830	-1.0	GD
10835	-0.6	NE
10840	0.2	NE
10845	-inf	ES
10850	-inf	GD
10855	-0.47	NE
10860	0.12	NE
10865	-0.45	GD
10870	-0.38	NE
10875	-0.41	GD
10880	0.09	NE
10890	0.47	GA
10895	0.85	GA
10900	0.54	GA
10905	0.84	GA
10910	0.61	NE
10915	0.36	GA
10920	0.91	GA
10925	1.18	GA
10930	0.76	GA
10935	0.49	GA
10940	-0.02	NE
10945	0.58	NE
10950	-0.38	GD
10955	0.16	NE
10960	-0.37	GD
10965	-0.88	GD
10970	0.51	NE
10975	0.46	NE
10980	-0.29	NE
10985	-0.93	GD
10990	-inf	GD
11000	-inf	GD
11005	0.42	NE
11010	0.38	GA
11015	0.9	GA
11020	0.55	GA
11025	-inf	GD
11030	-inf	ES
11035	-inf	ES
11040	-inf	GD
11045	-inf	ES
11050	-inf	ES
11055	-inf	GD
11060	-0.52	NE
11065	-inf	ES
11070	-inf	ES
11075	0.45	GA
11080	-1.14	GD
11085	-0.6	GD
11090	-0.93	GD
11095	-0.05	NE
11100	-0.58	NE
11105	0.19	NE
11110	0.37	NE
11115	0.01	NE
11120	0.22	NE
11125	-0.09	NE
11130	-0.28	NE
11135	-0.02	NE
11140	-0.31	NE
11145	-0.65	GD
11150	-0.5	NE
11155	-1.04	GD
11160	-0.34	GD
11165	-0.57	NE
11170	0.47	NE
11175	1.03	GA
11180	-0.03	NE
11185	-1.23	GD
11190	0.09	NE
11195	-0.12	NE
11200	-0.6	NE
11205	-0.0	NE
11210	0.23	NE
11215	-inf	ES
11220	0.71	NE
11225	-0.88	NE
11230	0.16	NE
11235	0.81	NE
11240	-0.37	NE
11245	-inf	ES
11250	0.18	NE
11255	-0.54	NE
11260	-0.23	NE
11265	-inf	ES
11270	-0.27	NE
11275	-0.18	NE
11280	0.29	NE
11285	-0.34	NE
11290	-0.16	NE
11295	-0.65	NE
11300	-0.42	NE
11305	-1.75	GD
11310	-inf	GD
11315	-inf	NE
11320	0.34	GA
11325	0.97	GA
11330	0.61	NE
11335	0.93	GA
11340	0.33	NE
11345	0.45	GA
11350	0.75	GA
11355	0.41	GA
11360	0.56	NE
11365	0.38	GA
11370	0.26	NE
11375	0.02	NE
11380	0.47	GA
11385	0.39	NE
11390	1.16	GA
11395	0.4	NE
11400	0.85	GA
11405	0.63	GA
11410	-0.01	NE
11415	-0.32	GD
11420	-0.33	GD
11425	-0.17	NE
11430	0.08	NE
11435	-0.31	GD
11440	-0.39	GD
11445	-0.23	GD
11450	0.03	NE
11455	-0.13	NE
11460	-0.2	NE
11465	-0.39	NE
11470	-0.01	NE
11475	-0.97	GD
11480	-0.42	NE
11485	0.37	NE
11490	-0.5	NE
11495	0.22	NE
11500	-0.15	NE
11505	0.09	NE
11510	0.01	NE
11515	-0.14	NE
11520	-0.44	NE
11525	0.01	NE
11530	0.48	GA
11535	-0.18	NE
11540	0.51	NE
11545	0.38	NE
11550	0.59	GA
11555	0.16	NE
11560	-inf	GD
11565	0.71	NE
11570	0.79	GA
11575	0.69	GA
11580	0.52	GA
11585	0.46	GA
11590	1.19	GA
11595	0.95	GA
11600	-0.01	NE
11605	-0.31	NE
11610	0.21	NE
11615	0.57	GA
11620	-0.26	NE
11625	0.83	GA
11630	0.6	NE
11635	0.57	GA
11640	0.77	GA
11645	0.4	GA
11650	0.74	GA
11655	0.28	NE
11660	0.38	NE
11665	0.21	NE
11670	0.66	NE
11675	0.56	NE
11680	0.24	NE
11685	0.4	NE

Feature	Log(RF)	Class
11690	0.54	GA
11695	0.52	GA
11700	0.83	GA
11705	0.38	NE
11710	0.79	GA
11715	0.32	NE
11720	-0.43	NE
11725	-0.29	NE
11730	-0.42	NE
11735	-0.26	NE
11740	-0.61	NE
11745	0.32	NE
11750	0.92	NE
11755	0.35	NE
11760	0.78	GA
11765	0.33	NE
11770	0.32	NE
11775	0.12	NE
11780	-0.05	NE
11785	0.26	NE
11790	0.7	GA
11795	0.48	GA
11800	-inf	NE
11805	0.65	GA
11810	0.4	GA
11815	-inf	ES
11820	0.34	NE
11825	0.59	NE
11830	-0.04	NE
11835	0.2	NE
11840	0.73	GA
11845	0.66	GA
11850	0.38	GA
11855	0.02	NE
11860	0.06	NE
11865	0.45	NE
11870	0.72	GA
11875	0.65	GA
11880	0.78	GA
11885	-1.62	NE
11890	-0.05	NE
11895	0.33	NE
11900	0.3	NE
11905	0.63	GA
11910	0.67	NE
11915	0.03	NE
11920	0.54	NE
11925	0.87	NE
11930	-0.38	NE
11935	0.5	NE
11940	0.15	NE
11945	0.08	NE
11950	-0.77	GD
11955	0.07	NE
11960	-0.37	GD
11965	-0.65	GD
11970	-1.05	NE
11975	-1.03	GD
11980	-0.34	NE
11985	-0.99	GD
11990	-0.61	GD
11995	-0.58	GD
12000	-0.91	GD
12005	-0.3	GD
12010	-0.33	GD
12015	-1.02	GD
12020	0.8	NE
12025	0.5	NE
12030	0.11	NE
12035	0.25	NE
12040	-0.01	NE
12045	-0.07	NE
12050	0.08	NE
12055	0.05	NE
12060	0.04	NE
12065	0.41	NE
12070	0.1	NE
12075	-0.57	NE
12080	-1.17	NE
12085	-0.46	NE
12090	0.1	NE
12095	0.54	GA
12100	-0.41	NE
12105	0.6	GA
12110	-1.05	GD
12115	0.75	NE
12120	-0.82	GD
12125	-0.21	NE
12130	-0.85	GD
12135	-0.45	GD
12140	0.1	NE
12145	-0.47	NE
12150	-0.64	NE
12155	-0.35	GD
12160	0.32	NE
12165	-0.77	GD
12170	-1.0	GD
12175	-1.32	GD
12180	-0.77	GD
12185	-0.29	GD
12190	-0.22	NE
12195	0.05	NE
12200	-0.36	NE
12205	-0.22	NE
12210	0.22	NE
12215	-0.49	GD
12220	-0.42	GD
12225	-0.85	GD
12230	-0.72	GD
12235	-0.66	GD
12240	-0.29	GD
12245	0.07	NE
12250	1.04	GA
12255	0.61	NE
12260	0.66	GA
12265	0.44	NE
12270	0.06	NE
12275	-0.32	NE
12280	0.11	NE
12285	0.19	NE
12290	-0.53	GD
12295	-0.62	GD
12300	-1.29	GD
12305	0.51	NE
12310	-0.1	NE
12315	-0.65	NE
12320	-0.15	NE
12325	0.43	NE
12330	-0.31	NE
12335	0.11	NE
12340	0.04	NE
12345	0.08	NE
12350	0.14	NE
12355	-1.35	GD
12360	-0.54	NE
12365	-0.12	NE
12370	-0.37	GD
12375	-0.51	GD
12380	-0.92	GD
12385	-0.65	GD
12390	0.32	NE
12395	-1.48	GD
12400	-0.33	GD
12405	-0.8	NE
12410	-0.87	NE
12420	0.6	GA
12425	0.57	NE
12430	0.59	NE
12435	-1.93	NE
12440	0.12	NE
12445	-0.1	NE
12450	-0.19	NE
12455	-0.57	GD
12460	0.37	NE
12465	-0.06	NE
12470	0.13	NE
12475	0.22	NE
12480	0.19	NE
12485	0.71	GA
12490	0.48	NE
12495	0.69	GA
12500	0.74	GA
12505	0.85	GA
12510	0.39	NE
12515	0.79	GA
12520	0.74	GA
12525	0.01	NE
12530	0.5	GA
12535	0.33	GA
12540	0.56	GA
12545	-0.34	GD
12550	0.22	NE
12555	-1.75	GD
12560	0.36	GA
12565	-0.05	NE
12570	-0.5	GD
12575	0.66	GA
12580	0.73	GA
12585	0.6	GA
12590	0.5	NE
12595	0.01	NE
12600	0.28	NE
12605	0.95	GA
12610	0.51	GA
12615	0.66	GA
12620	0.84	NE
12625	0.7	GA
12630	0.7	GA
12635	0.45	NE
12640	0.5	NE
12645	0.22	NE
12650	-0.23	NE
12655	0.2	NE
12660	0.28	NE
12665	0.36	GA
12670	-2.23	GD
12675	0.25	NE
12680	0.26	NE
12685	0.28	NE
12690	-0.67	NE
12695	-0.39	NE
12700	-inf	GD
12705	0.73	GA
12710	-inf	GD
12715	-1.27	GD
12720	0.39	NE
12725	0.01	NE
12730	-0.48	NE
12735	-inf	ES
12740	0.58	NE
12745	0.41	NE
12750	-0.22	NE
12755	-0.88	NE
12760	0.43	NE
12770	0.63	GA
12775	0.6	GA
12780	0.85	NE
12785	0.68	NE
12790	-0.1	NE
12795	-1.56	ES
12800	0.14	NE
12805	0.6	GA
12810	-0.27	NE
12815	0.68	NE
12820	1.0	GA
12825	1.39	GA
12830	0.79	GA
12835	0.6	NE
12840	1.03	GA
12845	0.66	GA
12850	0.23	NE
12855	0.65	GA
12860	0.47	NE
12865	0.02	NE
12870	0.02	NE
12875	0.66	GA
12880	0.09	NE
12885	0.06	NE
12890	0.05	NE
12895	-0.31	NE
12900	-0.03	NE
12905	0.83	NE
12910	0.31	NE
12915	-0.44	NE
12920	0.16	NE
12925	-0.98	GD
12930	0.18	NE
12935	0.32	NE
12940	0.28	NE
12945	-0.62	NE
12950	0.23	NE
12955	0.19	NE
12960	0.86	NE
12965	0.31	NE
12970	-0.1	NE
12975	0.74	NE
12980	0.03	NE
12985	-0.57	NE
12990	-0.14	NE
12995	0.47	NE
13000	0.14	NE
13005	-0.04	NE
13010	0.58	NE
13015	0.59	NE
13020	0.69	GA
13025	0.41	NE
13030	0.79	GA
13035	0.97	GA
13040	0.14	NE
13045	-0.6	NE
13050	-1.93	GD
13055	0.14	NE
13060	0.55	NE
13065	0.43	NE
13070	0.67	NE
13075	0.36	NE
13080	0.35	GA
13085	0.19	NE
13090	-0.15	NE
13095	-1.09	GD
13100	-0.32	GD
13105	-0.28	NE
13110	0.43	NE
13115	0.06	NE
13120	0.14	NE
13125	0.54	NE
13130	-0.03	NE
13135	0.41	NE
13140	0.54	NE
13145	-0.16	NE
13150	0.52	NE
13155	-0.07	NE
13160	0.55	NE
13165	0.38	NE
13170	0.42	GA
13180	0.38	NE
13185	0.11	NE
13190	0.52	NE
13195	0.24	NE
13200	0.37	NE
13205	0.22	NE
13210	-0.23	NE
13215	0.26	NE
13220	0.11	NE
13225	0.6	NE
13230	0.66	NE
13235	-0.15	NE
13240	0.17	NE
13245	0.3	NE
13250	0.52	NE
13255	0.14	NE
13260	0.64	GA
13265	0.9	GA
13270	0.64	GA
13275	0.52	NE
13280	0.01	NE
13285	-0.58	NE
13290	-0.84	GD
13295	-0.06	NE
13300	0.35	NE
13305	0.01	NE
13310	0.7	NE
13315	0.09	NE
13320	-0.13	NE
13325	0.4	NE
13330	-0.03	NE
13335	-0.0	NE
13340	0.13	NE
13345	-inf	NE
13350	-0.17	NE
13355	0.43	NE
13360	0.89	GA
13365	0.9	GA
13370	-0.6	NE
13375	-0.61	NE
13380	0.09	NE
13385	-0.09	NE
13390	-0.76	GD
13395	-0.1	NE
13400	0.01	NE
13405	-0.46	NE
13410	-0.34	GD
13415	-0.11	NE
13420	0.05	NE
13425	0.86	GA
13430	0.06	NE
13435	0.11	NE
13440	0.15	NE
13445	0.74	GA
13450	-0.64	GD
13455	-0.8	NE
13460	0.11	NE
13465	0.39	NE
13470	0.8	GA
13475	0.24	NE
13480	-0.03	NE
13485	0.28	NE
13490	-1.32	GD
13495	0.14	NE
13500	-0.89	NE
13505	0.57	NE
13510	0.67	NE
13515	-0.33	NE
13520	-0.43	NE
13525	-0.26	GD
13530	-0.83	GD
13535	-0.2	NE
13540	-0.65	GD
13545	-0.3	GD
13550	-1.45	GD
13555	-0.43	NE
13560	-0.43	GD
13565	0.79	GA
13570	0.38	NE
13575	-0.01	NE
13580	-0.5	NE
13585	-0.69	GD
13590	0.04	NE
13595	0.01	NE
13600	-1.02	GD
13605	-0.19	NE
13610	0.31	NE
13615	0.3	GA
13620	-0.12	NE

Feature	Log(RF)	Class
13625	0.21	NE
13630	0.21	NE
13635	0.68	NE
13640	1.25	GA
13645	0.73	GA
13650	-0.15	NE
13655	0.53	GA
13660	0.72	GA
13665	0.37	GA
13675	0.92	GA
13680	0.71	GA
13685	0.18	NE
13690	0.23	NE
13695	0.37	NE
13700	0.57	GA
13705	0.34	NE
13710	0.69	GA
13715	0.7	NE
13720	0.51	GA
13725	0.82	GA
13730	0.48	NE
13735	0.64	GA
13740	0.03	NE
13745	0.03	NE
13755	0.27	NE
13760	-0.28	GD
13765	-0.56	NE
13770	-0.1	NE
13775	-0.86	GD
13780	-0.76	GD
13785	-0.72	GD
13790	-0.36	GD
13795	-0.05	NE
13800	0.14	NE
13805	0.79	NE
13810	-0.12	NE
13815	0.12	NE
13820	-1.09	GD
13825	-1.15	NE
13830	-0.79	GD
13835	-0.45	NE
13840	-0.21	NE
13845	-0.68	NE
13850	-1.04	GD
13855	-1.11	GD
13860	-0.18	NE
13865	0.47	NE
13870	-0.53	NE
13875	-0.26	NE
13880	-0.12	NE
13885	-0.14	NE
13890	-0.64	GD
13895	-0.59	NE
13900	-0.53	NE
13905	-0.71	NE
13910	-0.8	NE
13915	-0.6	GD
13920	-0.37	GD
13925	-0.35	NE
13930	0.4	NE
13935	-0.15	NE
13940	-inf	NE
13945	0.05	NE
13950	-0.69	GD
13955	0.1	NE
13960	0.74	GA
13965	-0.19	NE
13970	0.43	NE
13975	0.32	NE
13980	-0.79	NE
13985	0.1	NE
13990	-0.48	NE
13995	0.11	NE
14000	-0.71	NE
14005	-0.07	NE
14010	-0.26	GD
14015	-0.23	GD
14020	0.31	NE
14025	-0.15	NE
14030	-0.45	NE
14035	-0.29	GD
14040	-0.08	NE
14045	0.11	NE
14050	-0.99	GD
14055	0.38	NE
14060	-0.6	GD
14065	-0.15	NE
14070	-0.93	GD
14075	-1.19	GD
14080	-1.67	GD
14085	-0.94	GD
14090	-1.97	ES
14095	-inf	ES
14100	0.68	NE
14105	-inf	NE
14110	-inf	ES
14115	0.06	NE
14120	0.73	GA
14125	0.23	NE
14130	-0.56	NE
14135	-0.08	NE
14140	-0.57	GD
14145	-0.07	NE
14150	-0.72	GD
14155	-1.97	GD
14160	-0.07	NE
14165	0.14	NE
14170	0.59	NE
14175	0.36	NE
14180	0.01	NE
14185	-0.16	NE
14190	-2.23	GD
14195	-0.72	GD
14200	-1.22	GD
14205	0.48	NE
14210	0.64	GA
14215	-0.77	GD
14220	-0.98	GD
14225	0.12	NE
14230	-0.33	GD
14235	0.98	GA
14245	0.45	GA
14250	0.53	NE
14255	-inf	ES
14260	-0.04	NE
14265	-0.41	GD
14270	0.26	NE
14275	0.44	GA
14280	0.32	NE
14285	-inf	GD
14290	-1.37	GD
14295	0.9	GA
14300	0.27	NE
14305	0.1	NE
14310	-0.05	NE
14315	0.51	NE
14320	1.27	GA
14325	-0.04	NE
14330	0.83	GA
14335	0.22	NE
14340	0.34	NE
14345	0.04	NE
14350	0.91	GA
14355	0.12	NE
14360	0.13	NE
14365	0.11	NE
14370	-inf	ES
14375	-inf	GD
14380	-0.84	GD
14385	0.5	NE
14390	-0.13	NE
14395	-1.32	GD
14400	-0.54	GD
14405	-0.87	GD
14410	-0.98	NE
14415	-0.26	NE
14420	-0.75	GD
14425	-1.25	NE
14430	-0.78	GD
14435	-0.67	GD
14440	0.66	GA
14445	0.57	NE
14450	-1.32	GD
14455	-0.55	GD
14460	-0.85	NE
14465	-0.4	GD
14470	-0.89	GD
14475	-0.49	GD
14480	-0.52	GD
14485	-0.62	NE
14490	-0.73	GD
14495	-0.81	GD
14500	0.25	NE
14505	-0.48	GD
14510	0.56	NE
14515	0.76	NE
14520	0.44	NE
14525	0.57	NE
14530	0.32	GA
14535	-1.17	NE
14540	-0.24	NE
14545	-inf	GD
14550	-inf	ES
14555	-0.23	NE
14560	0.59	GA
14565	0.47	GA
14570	0.77	NE
14575	-inf	GD
14580	0.24	NE
14585	-0.11	NE
14590	0.13	NE
14595	0.63	GA
14600	-0.98	NE
14605	-0.05	NE
14610	0.09	NE
14620	-0.14	NE
14625	0.14	NE
14630	0.09	NE
14635	-0.11	NE
14640	-0.19	NE
14645	-0.01	NE
14650	0.49	NE
14655	0.51	NE
14660	0.21	NE
14665	-inf	ES
14675	-inf	ES
14680	-inf	NE
14685	-0.65	NE
14690	-inf	GD
14695	-1.14	GD
14700	-inf	GD
14705	-inf	ES
14710	-inf	ES
14715	-inf	GD
14720	-inf	GD
14725	-inf	GD
14730	-inf	GD
14735	-inf	ES
14740	-inf	ES
14745	-inf	GD
14750	-inf	ES
14755	-inf	GD
14760	-0.34	NE
14765	0.4	NE
14770	0.48	NE
14775	-1.34	GD
14780	-0.93	GD
14785	-inf	ES
14790	0.08	NE
14795	-0.5	GD
14800	-0.66	GD
14805	-0.2	NE
14810	0.11	NE
14815	0.35	NE
14820	0.2	NE
14825	-inf	ES
14830	-0.4	NE
14835	-0.33	NE
14840	-0.39	NE
14845	0.07	NE
14850	0.54	GA
14855	-inf	GD
14860	0.53	NE
14865	0.31	NE
14870	-0.23	NE
14875	-0.27	NE
14880	-inf	ES
14885	-inf	ES
14890	-1.81	GD
14895	-inf	ES
14900	-1.97	ES
14905	-0.19	NE
14910	0.07	NE
14915	-0.22	NE
14920	-1.32	NE
14925	-1.66	GD
14930	-inf	ES
14935	-0.23	NE
14940	0.32	NE
14945	-1.64	NE
14950	-0.31	NE
14955	0.15	NE
14960	0.1	NE
14965	-0.76	GD
14970	-0.46	NE
14975	-0.64	GD
14980	-1.46	GD
14985	-0.06	NE
14990	-0.06	NE
14995	-0.39	NE
15000	-0.03	NE
15005	-1.31	GD
15010	-0.2	NE
15015	-inf	GD
15020	-2.23	GD
15025	-0.27	NE
15030	0.04	NE
15035	-inf	ES
15040	-0.05	NE
15045	0.41	NE
15050	-0.35	NE
15055	-inf	ES
15060	0.4	NE
15065	0.27	GA
15070	0.56	NE
15075	-inf	GD
15080	0.51	NE
15085	0.65	NE
15090	0.16	NE
15095	-0.38	GD
15100	-inf	GD
15105	-1.09	GD
15110	-0.33	NE
15115	0.91	NE
15120	-1.01	GD
15125	0.44	NE
15130	-1.05	GD
15135	-1.03	NE
15140	-0.01	NE
15145	-0.7	NE
15150	-1.06	GD
15155	-0.48	NE
15160	-inf	GD
15165	-0.88	GD
15170	-0.7	GD
15175	-0.18	GD
15180	-0.49	NE
15185	-0.31	NE
15190	-0.47	NE
15195	0.9	NE
15200	0.44	NE
15205	-0.65	GD
15210	-inf	ES
15215	-inf	GD
15220	-inf	GD
15225	-inf	ES
15230	-0.1	NE
15235	-0.12	NE
15240	-0.64	GD
15245	-0.27	GD
15250	-0.2	NE
15255	-0.59	GD
15260	-0.15	NE
15265	1.08	NE
15270	0.5	NE
15275	0.81	GA
15280	-0.22	NE
15285	-0.04	NE
15290	0.05	NE
15295	-0.17	NE
15300	0.26	NE
15305	0.32	NE
15310	0.13	NE
15315	-0.83	GD
15320	-0.31	GD
15325	-0.11	NE
15330	0.46	NE
15335	0.45	NE
15340	-0.64	NE
15345	-0.77	GD
15350	-0.09	NE
15355	0.15	NE
15360	-0.57	GD
15365	0.27	NE
15370	0.11	NE
15375	0.56	GA
15380	-0.53	GD
15385	0.35	GA
15390	0.03	NE
15395	0.36	NE
15400	-0.15	NE
15405	-0.55	GD
15410	-1.34	GD
15415	-1.16	GD
15420	0.29	NE
15425	-0.28	GD
15430	-0.06	NE
15435	0.13	NE
15440	-0.48	NE
15450	0.28	NE
15455	0.3	NE
15460	-0.13	NE
15465	-0.36	NE
15470	0.09	NE
15475	-0.36	GD
15480	-1.04	GD
15485	-0.06	NE
15490	-1.24	GD
15495	-0.17	NE
15500	-0.2	GD
15505	-0.99	GD
15510	-0.73	NE
15515	-0.35	GD
15520	-1.03	GD
15525	-1.09	GD
15530	-0.05	NE
15535	-0.24	NE
15540	0.32	NE
15545	0.1	NE
15550	-0.76	GD
15555	0.24	NE
15560	-0.22	NE
15565	-0.58	GD
15570	-1.32	GD

Feature	Log(RF)	Class
15575	-1.27	GD
15580	-0.51	GD
15585	-0.17	NE
15595	-0.04	NE
15600	-0.66	NE
15605	-0.98	GD
15610	-0.14	NE
15615	-1.22	GD
15620	-0.48	GD
15625	-1.15	GD
15630	-0.07	NE
15640	0.34	NE
15645	-0.72	GD
15650	-0.68	GD
15655	-0.7	GD
15660	0.25	NE
15665	0.31	NE
15670	0.19	NE
15675	0.13	NE
15680	0.28	NE
15685	0.47	NE
15690	-inf	GD
15695	-0.8	GD
15700	-0.33	GD
15705	-0.36	GD
15710	-0.08	NE
15715	0.53	NE
15720	0.53	GA
15725	-0.48	GD
15730	0.44	NE
15735	0.42	GA
15745	0.82	GA
15750	0.61	GA
15755	0.56	GA
15760	0.53	NE
15765	0.23	NE
15770	0.92	GA
15775	0.82	NE
15780	0.43	GA
15785	0.6	GA
15790	0.74	NE
15795	0.41	GA
15800	0.69	GA
15805	1.02	NE
15810	0.68	GA
15815	0.58	GA
15820	1.08	GA
15825	0.57	NE
15830	0.42	GA
15835	0.73	NE
15840	0.5	NE
15845	0.31	NE
15850	0.29	NE
15855	-0.45	NE
15860	-0.35	NE
15865	-0.13	NE
15870	-1.62	GD
15875	-0.22	NE
15880	-1.51	GD
15885	-1.09	GD
15890	-0.78	GD
15895	-0.83	GD
15900	-0.46	GD
15905	0.11	NE
15910	-0.91	GD
15915	-0.54	NE
15925	0.16	NE
15930	0.32	NE
15935	0.2	NE
15940	-0.42	GD
15945	-0.44	GD
15950	-inf	NE
15955	-0.27	NE
15960	-0.9	GD
15965	-inf	NE
15970	-1.27	GD
15975	-1.08	GD
15980	-1.28	GD
15985	-0.43	NE
15990	-1.15	GD
15995	-1.62	GD
16000	-0.42	GD
16005	-0.62	NE
16010	-0.19	NE
16015	-1.04	GD
16020	-0.15	NE
16025	0.23	NE
16030	-0.15	NE
16035	-1.32	GD
16040	-1.93	GD
16045	-1.23	GD
16050	-0.98	GD
16055	-0.2	NE
16060	-inf	NE
16065	-inf	GD
16070	-0.14	NE
16075	-inf	ES
16080	-0.37	NE
16085	-0.6	GD
16090	-0.36	NE
16095	-0.07	NE
16100	0.36	NE
16105	-inf	ES
16110	-0.78	NE
16115	-0.19	NE
16120	-0.8	NE
16125	-inf	ES
16130	-0.4	NE
16135	0.56	NE
16140	-0.87	GD
16145	-1.41	GD
16150	-0.95	GD
16155	0.15	NE
16160	-0.54	NE
16165	0.68	GA
16170	0.57	NE
16175	0.29	NE
16180	0.05	NE
16185	0.15	NE
16190	0.59	NE
16195	-0.88	NE
16200	1.06	NE
16205	-0.41	NE
16210	-1.93	NE
16215	-0.42	GD
16220	-0.94	GD
16225	-1.08	GD
16230	-1.67	GD
16235	-1.93	GD
16240	-0.65	GD
16245	-1.47	GD
16250	-0.71	GD
16255	-0.04	NE
16260	-0.74	NE
16265	-0.4	GD
16270	-1.34	GD
16275	-0.42	GD
16280	-0.86	GD
16285	-0.71	GD
16290	-0.75	GD
16295	-inf	GD
16300	-0.11	NE
16305	-0.3	GD
16310	-0.43	NE
16315	-0.6	NE
16320	-0.07	NE
16325	-0.82	GD
16330	-0.11	NE
16335	-1.47	GD
16340	-0.63	GD
16345	-0.67	NE
16350	-0.85	GD
16355	-0.83	GD
16360	0.05	NE
16365	-0.6	GD
16370	-0.39	GD
16375	-0.18	NE
16380	-0.11	NE
16385	-0.36	NE
16390	-1.24	GD
16395	-0.91	GD
16400	-0.91	GD
16405	-0.32	GD
16410	0.44	NE
16415	0.52	NE
16420	0.62	NE
16425	-0.12	NE
16430	-0.06	NE
16435	0.23	NE
16440	0.09	NE
16445	0.42	GA
16450	-0.16	NE
16455	0.43	GA
16460	-0.68	GD
16465	0.79	NE
16470	0.06	NE
16475	-0.59	GD
16480	-0.26	NE
16485	-0.1	NE
16490	-0.49	NE
16495	-0.02	NE
16500	-0.55	NE
16510	0.27	NE
16515	-1.66	GD
16520	0.15	NE
16525	0.26	NE
16530	0.79	GA
16535	0.13	NE
16540	-0.55	NE
16545	0.17	NE
16550	-0.08	NE
16555	0.4	GA
16565	-inf	GD
16570	0.74	NE
16580	-0.62	NE
16585	-0.63	GD
16590	-0.78	GD
16595	-0.65	GD
16600	0.14	NE
16605	-0.07	NE
16610	-0.09	NE
16615	-0.64	GD
16620	0.09	NE
16625	-0.57	GD
16630	-0.3	NE
16635	0.08	NE
16640	-0.54	NE
16645	-1.06	GD
16650	-0.44	NE
16655	-0.24	NE
16660	-inf	ES
16665	-0.58	GD
16670	-0.04	NE
16675	-0.21	NE
16680	-0.27	NE
16685	-0.45	GD
16690	-1.48	NE
16695	-0.19	NE
16700	-0.93	GD
16705	-0.47	NE
16710	-0.38	GD
16715	-0.07	NE
16720	-0.77	GD
16725	-inf	ES
16730	0.35	NE
16735	-0.25	GD
16740	-0.41	NE
16745	-0.37	NE
16750	-inf	ES
16755	-inf	GD
16765	-1.13	GD
16770	-1.54	GD
16775	-1.46	GD
16780	-0.2	NE
16785	-1.64	ES
16790	-inf	ES
16795	0.35	NE
16800	0.03	NE
16805	0.65	GA
16810	0.14	NE
16815	0.51	NE
16820	0.26	GA
16825	-inf	ES
16830	0.32	NE
16835	-inf	GD
16840	-inf	GD
16845	0.21	NE
16850	-inf	GD
16855	-1.93	GD
16860	0.49	NE
16870	-0.53	NE
16875	-inf	GD
16880	-0.11	NE
16885	0.0	NE
16890	0.44	NE
16895	0.79	NE
16900	0.29	NE
16905	0.43	NE
16910	0.49	NE
16915	-1.23	GD
16920	-0.72	GD
16925	-0.3	NE
16930	-0.31	GD
16935	-inf	ES
16940	0.59	NE
16945	-1.81	GD
16950	0.16	NE
16955	-0.38	NE
16960	-0.1	NE
16965	0.11	NE
16970	-inf	GD
16975	0.34	GA
16980	0.17	NE
16985	-0.71	NE
16990	-0.64	GD
16995	0.67	NE
17000	-1.3	GD
17005	-0.9	GD
17010	-1.06	GD
17015	-0.4	NE
17020	0.49	NE
17025	-0.75	GD
17030	-1.04	GD
17035	-0.44	GD
17040	0.05	NE
17045	0.31	NE
17050	0.36	NE
17055	-inf	GD
17060	-0.99	GD
17065	-0.36	NE
17070	-inf	GD
17075	0.29	NE
17080	0.36	NE
17085	-0.04	NE
17090	0.45	GA
17095	0.08	NE
17100	-inf	GD
17105	-0.03	NE
17110	0.61	GA
17115	-inf	NE
17120	0.27	NE
17125	0.21	NE
17130	0.22	NE
17135	-0.01	NE
17140	-0.01	NE
17145	0.29	GA
17150	-1.17	GD
17155	-0.46	NE
17160	-0.07	NE
17165	0.32	GA
17170	0.2	NE
17175	0.59	NE
17180	-0.09	NE
17185	-0.21	NE
17190	0.33	GA
17195	0.07	NE
17200	0.4	GA
17205	0.21	NE
17210	0.07	NE
17215	0.04	NE
17220	-0.55	NE
17230	-inf	ES
17235	0.52	NE
17240	-1.01	GD
17245	0.42	GA
17250	-0.57	NE
17255	-0.19	NE
17260	-1.12	GD
17265	-0.26	NE
17270	-0.52	GD
17275	-0.81	NE
17280	-inf	GD
17285	-0.56	GD
17290	-0.21	GD
17295	-0.04	NE
17305	-1.45	GD
17310	-0.15	NE
17315	-0.11	NE
17320	-0.28	NE
17325	0.26	NE
17330	-0.27	NE
17335	-0.49	GD
17340	-0.37	GD
17345	0.44	NE
17350	-0.07	NE
17355	0.74	GA
17360	-0.28	NE
17365	-0.93	GD
17370	-0.26	NE
17375	-1.06	GD
17380	-0.35	GD
17385	-0.44	GD
17390	0.0	NE
17395	-0.85	GD
17400	-0.07	NE
17405	-0.85	GD
17410	-0.52	GD
17415	0.67	NE
17420	0.39	NE
17425	-0.73	GD
17430	-0.45	GD
17435	-0.08	NE
17440	-0.25	GD
17445	-1.08	NE
17450	-0.98	GD
17455	-0.6	GD
17460	-1.04	GD
17465	-1.02	GD
17470	-0.62	GD
17475	-1.23	GD
17480	-0.36	GD
17485	-0.78	GD
17490	-1.45	GD
17495	-1.56	GD
17500	-0.9	GD
17505	-1.34	GD
17510	-1.31	GD
17515	-1.28	GD
17520	-0.16	NE
17525	-0.02	NE
17530	-1.56	NE
17535	-1.01	GD
17540	-0.59	GD
17545	-0.52	GD

Feature	Log(RF)	Class
17550	-1.97	GD
17555	-1.4	GD
17560	-inf	NE
17565	-0.83	GD
17570	-0.78	GD
17575	-inf	GD
17580	-0.77	GD
17585	-1.56	GD
17590	-0.13	NE
17595	-0.93	NE
17600	-0.67	GD
17605	-0.21	NE
17610	-inf	GD
17615	-0.6	GD
17620	-0.52	NE
17625	-0.49	NE
17630	0.37	NE
17635	-1.46	GD
17640	-0.29	NE
17645	-inf	GD
17650	0.16	NE
17655	-0.55	NE
17660	-0.26	NE
17665	0.63	NE
17670	0.23	NE
17675	-0.0	NE
17680	-0.05	NE
17685	0.11	NE
17690	0.1	NE
17695	-0.79	NE
17700	-1.15	GD
17705	-0.18	NE
17710	-0.29	NE
17715	-0.49	NE
17720	-0.37	GD
17725	-0.6	NE
17730	0.27	NE
17735	-0.4	NE
17740	0.72	NE
17745	0.71	GA
17750	0.27	NE
17755	-0.28	NE
17760	0.1	NE
17765	-inf	ES
17770	-inf	ES
17775	-0.21	NE
17780	-0.16	NE
17785	-0.19	NE
17790	-0.97	NE
17795	-0.0	NE
17800	0.02	NE
17805	0.25	NE
17810	-0.08	NE
17815	0.13	NE
17820	-0.02	NE
17825	-0.9	GD
17830	-0.05	NE
17835	-0.05	NE
17840	0.2	GA
17845	0.74	GA
17850	-0.22	NE
17855	0.43	GA
17860	0.44	GA
17865	-0.42	NE
17870	-inf	ES
17875	-inf	ES
17880	-inf	GD
17885	-0.17	NE
17890	-0.98	NE
17895	-0.84	NE
17900	-inf	ES
17905	-inf	ES
17910	-inf	ES
17915	-inf	ES
17920	-inf	GD
17925	-inf	GD
17930	-inf	GD
17935	-inf	GD
17940	-inf	ES
17945	0.11	NE
17950	-inf	NE
17955	-inf	ES
17960	-0.34	NE
17965	-inf	ES
17970	-0.14	NE
17975	-inf	ES
17980	-inf	GD
17985	-inf	ES
17990	-inf	ES
17995	-inf	ES
18000	-inf	GD
18005	0.11	NE
18010	0.02	NE
18015	0.61	NE
18020	1.0	NE
18025	0.73	NE
18030	0.2	NE
18035	-0.05	NE
18040	-0.27	NE
18045	0.66	GA
18055	0.22	NE
18060	0.42	NE
18065	0.71	NE
18070	0.44	NE
18075	0.34	NE
18080	0.09	NE
18085	0.63	NE
18090	0.58	GA
18095	0.59	GA
18100	0.14	NE
18105	0.4	NE
18110	0.54	NE
18115	0.53	GA
18120	0.58	GA
18125	0.34	NE
18130	0.94	NE
18135	0.45	GA
18140	0.73	GA
18145	0.66	GA
18150	0.27	GA
18155	0.69	GA
18160	0.46	NE
18165	0.09	NE
18170	-inf	GD
18175	0.29	NE
18180	0.44	GA
18185	0.09	NE
18190	0.28	NE
18195	0.23	NE
18200	0.54	NE
18205	-0.18	NE
18210	0.02	NE
18215	0.28	NE
18220	0.27	NE
18225	-0.6	NE
18230	-0.54	GD
18235	-0.21	NE
18240	0.45	NE
18245	-1.46	NE
18250	-0.5	GD
18255	-0.67	GD
18260	-0.78	GD
18265	-0.51	GD
18270	-0.76	GD
18275	-0.76	GD
18280	0.09	NE
18285	0.12	NE
18290	0.02	NE
18295	0.21	NE
18300	0.06	NE
18305	-0.44	GD
18310	-1.09	GD
18315	-0.56	NE
18320	-0.21	NE
18325	-0.17	NE
18330	-1.24	GD
18335	-0.01	NE
18340	0.46	NE
18345	-0.06	NE
18350	0.42	NE
18355	-0.36	GD
18360	-0.32	NE
18365	-0.28	NE
18370	-0.32	NE
18375	0.33	NE
18380	0.01	NE
18385	0.29	NE
18390	-0.23	NE
18395	-inf	ES
18400	-0.54	GD
18405	0.12	NE
18410	-0.32	NE
18415	-0.2	NE
18420	-0.39	GD
18425	-0.4	NE
18430	-0.7	GD
18435	-0.87	GD
18440	-0.73	GD
18445	-0.04	NE
18450	-0.01	NE
18455	-1.03	GD
18460	-0.37	GD
18465	-0.67	GD
18470	-0.23	NE
18475	0.33	NE
18480	-0.14	NE
18485	0.64	NE
18490	0.79	GA
18495	-1.33	GD
18500	-0.54	NE
18505	-inf	GD
18510	0.19	NE
18515	-0.47	GD
18520	-0.3	NE
18525	-0.95	GD
18530	-0.75	GD
18535	-0.6	GD
18540	-0.22	NE
18545	0.04	NE
18550	-0.62	NE
18555	0.45	GA
18560	0.19	NE
18565	0.15	NE
18570	-inf	GD
18575	-inf	NE
18580	-inf	GD
18585	0.45	NE
18590	0.44	NE
18595	0.43	NE
18600	0.18	NE
18605	-0.64	NE
18610	-0.56	GD
18615	-0.57	GD
18620	-1.16	GD
18625	0.39	NE
18630	-0.13	NE
18635	0.16	NE
18640	-0.03	NE
18645	0.18	NE
18650	-0.45	NE
18655	0.36	NE
18660	0.42	GA
18665	0.36	NE
18670	0.38	NE
18675	-inf	ES
18680	-0.13	NE
18685	-0.91	GD
18690	-inf	ES
18695	0.24	NE
18700	-0.6	NE
18705	0.1	NE
18710	-0.6	GD
18715	-0.55	NE
18720	-0.48	GD
18725	-0.65	GD
18730	-1.29	GD
18735	0.01	NE
18740	-0.89	GD
18745	-1.06	NE
18750	-0.65	GD
18755	-0.91	GD
18760	-0.74	GD
18765	-0.78	GD
18770	0.14	NE
18775	-0.59	GD
18780	-0.58	GD
18785	-0.52	NE
18790	-1.06	GD
18795	-0.18	NE
18800	-0.84	GD
18805	-0.68	GD
18810	-0.21	NE
18815	0.11	NE
18820	-0.75	GD
18825	-0.03	NE
18830	0.64	NE
18835	0.38	NE
18840	1.12	GA
18845	0.59	NE
18850	-0.22	NE
18855	0.04	NE
18860	0.22	NE
18865	0.1	NE
18870	-inf	ES
18875	0.45	GA
18880	0.44	NE
18885	0.12	NE
18890	-0.1	NE
18895	-0.49	GD
18900	-0.12	NE
18905	0.44	NE
18910	-0.23	NE
18915	-0.4	NE
18925	0.18	NE
18930	0.38	NE
18935	-0.1	NE
18945	0.26	NE
18950	0.19	NE
18955	0.04	NE
18960	0.1	NE
18965	-0.17	NE
18970	0.4	GA
18975	0.16	NE
18980	1.06	GA
18985	0.1	NE
18990	0.85	GA
18995	-0.78	GD
19000	-0.48	NE
19005	0.19	NE
19010	-0.49	GD
19015	-0.24	NE
19020	-0.72	GD
19025	-0.29	NE
19030	-0.88	GD
19035	-0.17	NE
19040	-0.45	GD
19045	-0.89	GD
19050	-1.43	GD
19055	-0.46	GD
19060	0.52	NE
19065	-0.31	GD
19070	-0.21	GD
19075	-0.2	GD
19080	-0.52	GD
19085	-0.28	NE
19090	0.4	NE
19095	-1.49	NE
19100	-0.08	NE
19105	0.07	NE
19110	0.24	NE
19115	-1.64	GD
19120	-1.62	GD
19125	-0.47	NE
19130	-0.33	GD
19135	0.16	NE
19140	-0.28	NE
19145	-0.71	GD
19150	-0.03	NE
19155	0.05	NE
19160	-0.65	NE
19165	-0.59	NE
19170	0.57	NE
19175	-0.51	NE
19180	-0.43	NE
19185	0.45	NE
19190	-0.08	NE
19195	-0.11	NE
19200	-0.64	GD
19205	0.62	NE
19210	0.32	NE
19215	0.38	GA
19220	0.42	NE
19225	-0.91	GD
19230	-inf	GD
19235	0.15	NE
19240	-0.45	NE
19245	-1.59	GD
19250	-0.5	NE
19255	-0.34	GD
19260	-0.71	GD
19265	0.22	NE
19270	-0.93	GD
19275	-inf	ES
19280	-1.67	GD
19285	-0.43	NE
19290	-1.32	GD
19295	-0.27	NE
19300	-inf	ES
19305	-inf	ES
19310	-inf	ES
19315	0.4	NE
19320	0.74	GA
19325	-inf	ES
19330	0.58	GA
19335	0.45	NE
19340	0.1	NE
19350	0.18	NE
19355	-0.76	NE
19360	0.16	NE
19365	-0.28	NE
19370	0.4	NE
19375	0.24	NE
19380	0.45	NE
19385	0.09	NE
19390	1.28	GA
19395	0.35	NE
19400	-0.36	NE
19405	-0.0	NE
19410	0.57	NE
19415	0.19	NE
19420	0.64	NE
19425	0.4	NE
19430	-0.62	NE
19435	0.31	NE
19440	-0.17	NE
19445	-1.32	NE
19450	-0.62	GD
19455	0.4	NE
19460	0.05	NE
19465	0.14	NE
19470	0.28	NE
19475	0.4	NE
19480	0.09	NE
19485	0.11	NE

Feature	Log(RF)	Class
19490	-inf	GD
19495	0.1	NE
19500	0.07	NE
19505	-0.34	NE
19510	0.12	NE
19515	0.11	NE
19520	-0.76	GD
19525	0.33	NE
19530	0.29	NE
19535	-0.19	GD
19540	-0.77	GD
19545	-0.89	GD
19550	-0.24	GD
19555	-0.67	GD
19560	-0.24	GD
19565	-0.48	GD
19570	-inf	GD
19575	-inf	GD
19580	-inf	ES
19585	-0.48	GD
19590	0.42	NE
19595	0.48	GA
19600	0.31	NE
19605	0.76	GA
19610	-inf	NE
19615	-inf	GD
19620	-inf	ES
19625	0.31	NE
19630	0.11	NE
19635	0.28	NE
19640	-1.12	NE
19645	-inf	GD
19650	0.47	GA
19655	0.6	NE
19660	-inf	GD
19665	0.57	NE
19670	0.91	GA
19675	0.88	GA
19680	-inf	ES
19685	0.3	NE
19690	0.58	GA
19700	-1.34	GD
19705	-0.73	GD
19710	-0.53	NE
19715	-0.02	NE
19720	0.51	NE
19725	-0.53	NE
19730	-0.46	GD
19735	-0.71	NE
19740	-1.67	GD
19745	-0.21	NE
19755	-0.26	NE
19760	-0.73	GD
19765	-0.93	GD
19770	-1.04	GD
19775	-0.74	GD
19780	-0.72	GD
19785	-0.83	GD
19790	-0.49	NE
19795	-0.37	NE
19800	-0.61	NE
19805	0.55	NE
19810	-0.07	NE
19815	0.05	NE
19820	0.23	NE
19825	0.15	NE
19830	0.15	NE
19835	0.07	NE
19840	0.09	NE
19845	0.09	NE
19850	-0.84	NE
19855	-1.24	GD
19860	-0.85	NE
19865	-0.2	NE
19870	-0.99	NE
19875	-0.26	GD
19880	-0.31	NE
19885	0.45	NE
19890	-0.42	GD
19895	-0.55	NE
19900	-inf	GD
19905	-0.32	NE
19910	0.35	GA
19915	0.56	NE
19920	0.34	GA
19925	0.02	NE
19930	0.31	NE
19935	-inf	GD
19940	0.37	GA
19945	0.47	NE
19950	-0.24	NE
19955	-inf	GD
19960	0.27	NE
19965	0.8	NE
19970	0.04	NE
19975	-0.08	NE
19980	-0.29	NE
19985	-0.04	NE
19990	0.21	NE
19995	-inf	ES
20000	0.39	GA
20005	-0.83	NE
20010	0.02	NE
20015	-0.78	GD
20020	-1.26	GD
20025	-1.12	GD
20030	-0.5	GD
20035	-0.44	GD
20040	-0.39	GD
20045	-1.08	GD
20050	-1.02	GD
20055	-0.35	NE
20060	0.43	GA
20065	-inf	ES
20070	-inf	GD
20075	-1.02	NE
20080	-0.28	NE
20085	0.51	NE
20090	0.68	NE
20095	-inf	GD
20100	-inf	GD
20105	0.04	NE
20110	-0.2	NE
20115	-0.74	GD
20120	-1.26	NE
20125	-1.21	GD
20130	-0.39	NE
20135	-0.9	GD
20140	0.06	NE
20145	-0.08	NE
20150	-0.07	NE
20155	-0.29	NE
20160	-0.02	NE
20165	0.21	NE
20170	-0.21	NE
20175	-0.75	GD
20180	-0.98	GD
20185	-0.28	NE
20190	-inf	ES
20195	-0.64	GD
20200	-1.03	GD
20205	-0.86	GD
20210	0.1	NE
20215	0.18	NE
20220	0.24	NE
20225	-0.11	NE
20230	-0.3	NE
20235	-0.02	NE
20240	-1.41	GD
20245	-0.19	GD
20250	-inf	GD
20255	-0.14	NE
20260	-0.15	NE
20265	-0.65	NE
20270	0.2	NE
20275	0.39	NE
20280	0.12	NE
20285	0.79	NE
20290	0.06	NE
20295	-0.43	GD
20300	-0.26	NE
20305	-0.68	NE
20310	-0.32	NE
20315	-0.77	GD
20320	0.05	NE
20325	-0.46	GD
20330	-0.49	NE
20335	-inf	GD
20340	-0.31	NE
20345	-0.28	NE
20350	-0.03	NE
20355	-0.33	GD
20360	-inf	GD
20365	0.43	NE
20370	-0.6	NE
20375	-0.62	NE
20380	-0.03	NE
20385	-inf	ES
20390	-1.26	NE
20400	0.3	NE
20410	0.82	NE
20415	0.73	NE
20420	0.31	NE
20425	-0.4	NE
20430	-0.38	NE
20435	0.44	NE
20440	-0.12	NE
20445	0.21	NE
20450	0.69	NE
20455	-0.0	NE
20460	0.29	NE
20465	1.07	GA
20470	0.11	NE
20475	0.97	NE
20480	0.3	NE
20485	0.81	NE
20490	1.14	GA
20495	0.13	NE
20500	-0.17	NE
20505	0.24	NE
20510	0.47	NE
20515	0.14	NE
20520	0.24	NE
20525	-0.55	NE
20530	-0.18	NE
20535	-1.41	GD
20540	-0.26	NE
20545	-0.72	GD
20550	0.15	NE
20555	0.08	NE
20560	-0.53	NE
20565	-0.12	NE
20570	0.18	NE
20575	0.96	NE
20580	-0.44	GD
20590	-0.48	NE
20595	-0.59	GD
20600	0.6	GA
20605	0.7	GA
20610	-0.33	NE
20615	0.44	GA
20620	-0.47	NE
20625	0.02	NE
20630	-0.46	NE
20635	-0.76	GD
20640	-0.67	GD
20645	-1.93	GD
20650	-0.46	GD
20655	0.21	NE
20660	-0.07	NE
20665	-0.12	NE
20670	-0.53	NE
20675	-0.15	NE
20680	-0.06	NE
20685	-0.94	GD
20690	-0.67	GD
20695	-0.7	GD
20700	-0.54	NE
20705	0.3	NE
20710	-0.43	GD
20715	0.74	GA
20720	0.64	GA
20725	0.92	GA
20735	0.16	NE
20740	0.22	NE
20745	-0.31	NE
20750	-0.61	NE
20755	-0.65	GD
20760	-0.85	GD
20765	-0.2	NE
20770	-0.56	GD
20775	0.22	NE
20780	-0.01	NE
20785	0.18	NE
20790	-0.38	NE
20795	-0.55	GD
20800	-1.1	GD
20805	-0.58	GD
20810	-1.13	GD
20815	0.44	NE
20820	-0.32	NE
20825	0.03	NE
20830	-0.35	NE
20835	0.05	NE
20840	0.23	NE
20845	-0.21	GD
20850	-0.21	NE
20855	0.1	NE
20860	-0.13	NE
20865	-0.32	NE
20870	-0.17	NE
20875	-0.02	NE
20880	-0.05	NE
20885	0.79	NE
20890	-0.95	NE
20895	-0.8	GD
20900	-0.18	NE
20905	0.26	NE
20910	-0.59	GD
20915	-0.7	GD
20920	-1.93	NE
20925	-0.84	GD
20930	-0.79	GD
20935	-0.78	NE
20940	-1.15	GD
20945	0.1	NE
20950	-0.29	NE
20955	-0.74	NE
20960	0.14	NE
20965	-0.08	NE
20970	-1.12	GD
20975	-1.05	GD
20980	-1.06	GD
20985	-0.67	GD
20990	-1.45	GD
21000	-0.78	GD
21005	-0.26	NE
21010	-0.89	GD
21015	-1.08	GD
21020	-1.25	GD
21025	-2.27	GD
21030	-0.85	GD
21035	-1.22	GD
21040	-0.68	GD
21045	-1.08	GD
21050	-0.9	GD
21055	-0.72	GD
21060	-1.32	GD
21065	-1.95	GD
21070	-0.33	NE
21075	-1.64	GD
21080	-0.41	NE
21085	-0.2	NE
21090	-0.55	GD
21095	-1.06	GD
21100	-0.55	GD
21105	-0.93	GD
21110	-1.81	GD
21115	0.07	NE
21120	0.39	NE
21125	0.14	NE
21130	-0.35	GD
21135	-0.9	GD
21140	-1.25	GD
21145	0.09	NE
21150	-0.25	GD
21155	-0.24	NE
21160	0.06	NE
21165	-0.02	NE
21170	-0.15	NE
21175	0.11	NE
21180	-0.13	NE
21185	-0.55	NE
21190	0.26	NE
21195	0.33	NE
21200	0.17	NE
21205	0.14	NE
21210	-0.0	NE
21215	0.62	NE
21220	0.15	NE
21225	-0.02	NE
21230	0.37	NE
21235	-0.05	NE
21240	0.03	NE
21245	0.7	GA
21250	0.06	NE
21255	-0.24	NE
21260	0.8	GA
21265	-inf	GD
21270	-inf	NE
21275	-inf	GD
21280	-inf	ES
21285	0.37	NE
21290	0.57	GA
21295	-0.12	NE
21300	0.29	NE
21305	0.6	GA
21310	0.21	NE
21315	-inf	ES
21320	0.55	GA
21325	0.7	NE
21330	0.53	GA
21335	-0.95	GD
21340	-0.83	GD
21345	0.18	GA
21350	0.25	GA
21355	0.38	GA
21360	0.33	GA
21365	0.52	GA
21370	0.42	GA
21375	-1.81	GD
21380	0.11	NE
21385	-0.71	NE
21390	0.81	GA
21395	-inf	NE
21400	0.39	GA
21405	0.68	NE
21410	-inf	ES
21415	-inf	NE
21420	-inf	ES
21425	-inf	ES
21430	-0.54	NE
21435	-1.37	GD
21440	-inf	ES

Feature	Log(RF)	Class
21445	0.58	NE
21450	-0.35	NE
21455	0.0	NE
21460	0.49	NE
21465	-0.48	NE
21470	0.02	NE
21475	0.15	NE
21480	0.17	NE
21485	-inf	GD
21495	0.33	NE
21500	-0.1	NE
21505	-0.48	GD
21510	-0.77	NE
21515	-0.02	NE
21520	0.45	GA
21525	0.59	NE
21530	0.28	NE
21535	-0.72	GD
21540	-inf	ES
21545	0.38	NE
21550	-0.52	GD
21555	-inf	ES
21560	0.64	NE
21565	-0.38	NE
21570	-0.31	NE
21575	-1.19	GD
21580	-inf	GD
21585	0.24	NE
21590	-1.04	GD
21595	0.1	NE
21600	-0.15	NE
21605	0.22	NE
21610	-0.67	GD
21615	0.49	NE
21620	-0.45	GD
21625	-0.38	NE
21630	-0.28	NE
21640	-0.49	GD
21645	-1.1	NE
21650	1.24	NE
21655	0.3	NE
21660	0.04	NE
21670	-inf	GD
21675	0.02	NE
21680	-0.11	NE
21685	0.04	NE
21690	0.59	NE
21695	-0.11	NE
21700	-0.1	NE
21705	-0.81	NE
21710	-0.35	GD
21715	0.65	GA
21720	0.51	GA
21725	0.1	NE
21730	-0.5	NE
21735	-0.22	NE
21740	0.38	NE
21745	0.76	NE
21750	0.18	NE
21755	0.04	NE
21760	-0.89	NE
21765	-0.42	GD
21770	-0.43	NE
21775	-0.41	GD
21780	-1.62	GD
21785	-0.59	GD
21790	-0.95	GD
21795	-0.23	GD
21800	-0.98	GD
21805	-0.18	NE
21810	-0.08	NE
21815	0.57	NE
21820	0.07	NE
21825	0.77	GA
21830	-0.15	NE
21835	-0.37	NE
21840	-inf	NE
21845	-0.69	NE
21850	-0.57	GD
21855	-0.12	NE
21860	-0.62	NE
21865	-0.13	NE
21870	-0.22	NE
21875	-0.26	NE
21880	-0.09	NE
21885	0.12	NE
21890	-0.14	NE
21895	0.26	NE
21900	-0.09	NE
21905	-0.39	NE
21910	0.17	NE
21915	-0.68	NE
21920	-0.1	NE
21925	-0.17	NE
21930	-0.47	GD
21935	-1.46	GD

Feature	Log(RF)	Class
21940	-0.61	GD
21945	-0.35	NE
21950	0.41	NE
21955	0.51	GA
21960	0.29	NE
21965	-0.22	GD
21970	-1.23	GD
21975	-0.8	GD
21980	-1.29	GD
21985	-0.66	GD
21990	-0.83	GD
21995	-0.65	GD
22000	-0.66	GD
22005	-0.46	NE
22010	0.5	NE
22015	-0.37	NE
22020	-0.23	NE
22025	-0.71	NE
22030	-0.3	GD
22035	-0.09	NE
22040	-0.59	GD
22045	-1.04	NE
22050	-0.81	GD
22055	0.2	NE
22060	-0.11	NE
22065	-1.0	GD
22070	0.26	NE
22075	-0.19	NE
22080	0.04	NE
22085	-0.66	GD
22090	-0.81	GD
22095	-0.6	GD
22100	-0.31	GD
22105	-0.95	GD
22110	-0.65	NE
22115	-0.69	NE
22120	-0.89	GD
22125	-0.89	GD
22130	-0.08	NE
22135	0.17	NE
22140	0.2	NE
22145	0.24	NE
22150	0.12	NE
22155	0.31	NE
22160	0.36	NE
22165	0.59	GA
22170	-0.73	GD
22175	-0.1	NE
22180	-0.26	NE
22185	0.42	NE
22190	-0.24	NE
22195	0.36	NE
22200	-0.17	NE
22205	0.2	NE
22210	-0.22	NE
22215	-0.49	NE
22220	-inf	ES
22230	-inf	GD
22235	-inf	ES
22240	0.27	NE
22245	0.98	GA
22250	-0.32	NE
22255	-inf	ES
22260	-1.56	NE
22265	-0.48	NE
22270	-0.25	NE
22275	0.6	NE
22280	-0.51	NE
22285	0.14	NE
22290	-inf	GD
22295	-0.4	NE
22300	0.4	NE
22305	-inf	ES
22310	-1.05	NE
22315	-inf	ES
22320	-0.24	NE
22325	-inf	GD
22330	-inf	ES
22335	-1.93	NE
22340	-0.74	GD
22345	0.79	NE
22350	-2.11	GD
22355	-1.93	GD
22360	-inf	GD
22365	-inf	GD
22370	-0.25	NE
22375	0.39	NE
22380	0.93	GA
22385	0.0	NE
22390	-0.32	GD
22395	-0.77	GD
22400	-0.41	GD
22405	-0.25	NE
22410	0.16	NE
22415	0.8	NE
22420	-1.31	GD

Feature	Log(RF)	Class
22425	-inf	GD
22430	0.16	NE
22435	-1.35	GD
22440	-inf	ES
22445	0.35	NE
22450	-1.97	GD
22455	-0.23	NE
22460	0.03	NE
22465	-inf	ES
22470	-inf	ES
22475	-inf	ES
22480	-inf	GD
22485	-0.04	NE
22490	-inf	ES
22495	-0.16	NE
22500	-0.69	NE
22505	-0.15	NE
22510	0.48	NE
22515	0.49	NE
22525	0.15	NE
22530	-0.33	NE
22535	-0.59	NE
22540	-1.56	NE
22545	0.35	NE
22550	0.23	GA
22555	-1.1	GD
22560	0.1	NE
22565	-0.4	GD
22570	-0.36	NE
22575	-0.39	GD
22580	-0.6	GD
22585	0.01	NE
22590	-0.19	NE
22595	0.49	NE
22600	-inf	GD
22605	-1.81	GD
22610	-1.93	GD
22615	-0.03	NE
22620	-1.93	GD
22625	-0.06	NE
22630	0.21	NE
22635	0.32	NE
22640	0.29	GA
22645	1.0	GA
22650	0.54	NE
22655	-0.86	GD
22660	-1.09	GD
22665	-0.09	NE
22670	0.51	GA
22675	-0.19	NE
22680	0.2	NE
22685	-inf	ES
22690	0.19	NE
22695	-0.33	GD
22700	0.2	NE
22705	0.38	NE
22710	0.68	NE
22715	0.02	NE
22720	0.26	NE
22725	-0.25	NE
22730	-0.09	NE
22735	-0.08	NE
22740	0.07	NE
22745	-1.64	GD
22750	-0.11	NE
22755	-1.09	GD
22760	-0.72	GD
22765	0.34	NE
22770	-0.84	NE
22775	-0.32	NE
22780	-1.64	NE
22785	-0.47	NE
22790	0.46	NE
22795	0.35	NE
22800	-0.25	GD
22805	-0.35	NE
22810	0.78	NE
22815	0.41	NE
22825	-1.81	GD
22830	-0.58	NE
22835	0.7	NE
22840	0.49	GA
22845	-0.97	NE
22850	0.21	NE
22855	0.52	NE
22860	0.18	NE
22865	0.62	NE
22870	-0.78	NE
22875	-1.09	NE
22880	-0.64	NE
22885	-1.26	GD
22890	-0.71	GD
22895	-0.65	GD
22900	-1.54	GD
22905	-0.17	NE
22910	-1.1	GD

Feature	Log(RF)	Class
22915	-0.15	NE
22920	-0.19	NE
22925	0.51	NE
22930	0.24	NE
22935	0.11	NE
22940	-0.18	NE
22945	-inf	NE
22950	-inf	ES
22955	-0.48	GD
22960	0.58	GA
22965	0.59	NE
22970	-inf	ES
22975	-inf	ES
22980	0.35	NE
22985	0.6	NE
22990	-0.41	NE
22995	-0.46	NE
23000	0.03	NE
23005	0.59	GA
23010	-inf	ES
23015	-inf	ES
23020	-0.56	NE
23025	-inf	GD
23030	-0.21	NE
23035	0.64	NE
23040	0.32	NE
23045	-1.93	GD
23050	-0.17	NE
23055	0.16	NE
23060	-0.32	NE
23065	-0.04	NE
23070	0.62	GA
23075	-0.0	NE
23080	0.07	NE
23085	0.13	NE
23090	-0.31	NE
23095	0.47	NE
23100	-0.16	NE
23105	-1.05	GD
23110	-0.0	NE
23115	0.5	NE
23120	0.12	NE
23125	-0.24	NE
23130	0.03	NE
23135	0.2	NE
23140	0.14	NE
23145	-inf	ES
23150	0.06	NE
23155	0.48	NE
23160	-0.84	NE
23165	-0.43	NE
23170	0.22	NE
23175	-0.06	NE
23180	0.14	NE
23190	-0.21	NE
23195	-0.27	NE
23200	0.45	NE
23205	0.01	NE
23210	0.22	NE
23215	-0.21	NE
23220	0.16	NE
23225	0.23	NE
23230	-0.26	NE
23235	0.11	NE
23240	-0.14	NE
23245	-0.22	NE
23250	-0.32	NE
23255	-inf	NE
23260	0.14	NE
23265	-0.17	NE
23270	-0.69	GD
23275	-0.77	NE
23280	-0.76	NE
23285	-1.04	GD
23290	-0.5	NE
23300	-0.18	NE
23305	-0.33	GD
23310	-1.13	GD
23315	-0.07	NE
23320	-0.55	GD
23325	-1.97	ES
23330	-0.4	GD
23335	-0.33	NE
23345	-0.07	NE
23350	-0.44	GD
23355	-0.97	GD
23360	-0.87	GD
23365	-0.48	NE
23370	-1.27	NE
23375	0.12	NE
23380	-0.06	NE
23385	-0.13	NE
23390	0.39	NE
23395	0.18	NE
23400	-0.12	NE
23405	-0.07	NE

Feature	Log(RF)	Class
23410	-0.26	NE
23415	0.12	NE
23420	-0.03	NE
23425	0.48	NE
23430	0.02	NE
23435	0.32	NE
23440	0.14	NE
23445	-0.05	NE
23450	-0.5	NE
23455	-0.23	NE
23460	0.27	NE
23465	0.58	NE
23470	0.63	NE
23475	0.37	NE
23480	-0.0	NE
23485	-0.07	NE
23490	0.04	NE
23495	-0.71	GD
23500	0.37	NE
23505	-inf	GD
23515	-0.42	NE
23520	-2.23	GD
23525	0.33	NE
23530	0.52	NE
23535	-1.08	NE
23540	-0.66	NE
23545	-0.71	GD
23555	-0.1	NE
23560	0.25	NE
23565	-0.03	NE
23570	0.18	NE
23575	-0.19	NE
23580	-0.45	GD
23585	-0.35	GD
23590	-0.52	NE
23595	-0.73	GD
23600	0.73	NE
23605	-0.96	GD
23610	0.15	NE
23615	0.22	NE
23620	-0.08	NE
23625	0.65	NE
23630	0.35	NE
23635	0.0	NE
23640	-inf	GD
23645	-inf	NE
23650	0.48	NE
23655	0.75	GA
23660	0.5	NE
23665	0.49	GA
23670	0.05	NE
23675	0.56	NE
23680	-0.83	GD
23685	-inf	GD
23690	0.29	NE
23695	-0.86	GD
23700	0.21	NE
23705	-inf	GD
23710	-inf	ES
23715	-inf	ES
23720	-inf	GD
23725	0.51	NE
23730	-1.4	GD
23735	0.29	NE
23740	-0.2	NE
23745	-inf	ES
23750	-inf	ES
23755	-inf	ES
23760	-inf	ES
23765	0.19	NE
23770	-0.33	NE
23775	0.34	NE
23780	-inf	ES
23785	-inf	ES
23790	0.38	NE
23795	0.18	NE
23800	0.25	NE
23805	-0.09	NE
23810	-inf	ES
23815	-inf	ES
23820	-inf	ES
23825	-0.05	NE
23830	0.62	GA
23835	-inf	ES
23840	-inf	ES
23845	-inf	ES
23850	-0.45	NE
23855	-inf	ES
23860	-inf	ES
23865	-2.23	GD
23870	0.48	NE
23875	0.35	NE
23880	0.53	GA
23885	0.54	NE
23890	0.06	NE
23895	-0.42	GD

Feature	Log(RF)	Class
23900	-1.45	GD
23905	0.03	NE
23910	0.31	NE
23915	0.45	NE
23920	-1.81	NE
23925	0.57	NE
23930	-inf	ES
23935	1.13	GA
23940	0.9	GA
23950	-0.06	NE
23955	0.06	NE
23960	0.21	NE
23965	0.51	NE
23970	0.21	NE
23975	0.8	GA
23980	-0.81	GD
23985	0.01	NE
23990	0.44	GA
23995	0.83	NE
24000	-0.36	NE
24005	-1.01	NE
24010	1.28	NE
24015	-1.45	NE
24020	0.69	GA
24025	0.46	NE
24030	0.6	NE
24035	0.17	NE
24040	0.63	GA
24045	0.31	NE
24050	-inf	ES
24055	-inf	ES
24060	0.18	NE
24065	-0.58	NE
24070	-0.68	NE
24075	0.12	NE
24080	-0.02	NE
24085	0.1	NE
24090	0.44	NE
24095	0.67	GA
24100	0.01	NE
24105	0.58	NE
24110	0.98	NE
24115	0.86	GA
24120	0.77	GA
24125	0.99	GA
24130	0.49	GA
24135	0.01	NE
24140	-0.29	NE
24145	-0.66	NE
24150	0.29	NE
24155	-0.4	GD
24160	-0.74	GD
24165	-0.34	NE
24170	-0.68	NE
24175	0.23	NE
24180	-1.76	GD
24185	0.62	NE
24190	-0.65	GD
24195	-0.4	GD
24200	0.17	NE
24205	-0.22	NE
24210	0.22	NE
24215	-0.88	NE
24220	-0.2	NE
24225	-1.64	GD
24230	0.5	NE
24235	0.63	GA
24240	0.98	NE
24245	0.76	GA
24250	-0.57	NE
24255	0.5	NE
24260	0.32	NE
24265	-1.62	GD
24270	-0.22	NE
24275	-0.09	NE
24280	-0.35	NE
24285	0.1	NE
24290	0.42	NE
24300	0.44	NE
24305	0.52	NE
24310	-0.06	NE
24315	-0.19	NE
24320	-0.22	NE
24325	-0.4	NE
24330	-0.29	GD
24335	-0.46	NE
24340	-0.35	GD
24345	-0.89	GD
24350	-1.02	GD
24355	-0.3	GD
24360	-0.41	NE
24365	-0.4	GD
24370	-0.81	GD
24375	0.11	NE
24380	-0.4	GD
24385	-0.33	NE

Feature	Log(RF)	Class
24390	-0.61	NE
24395	-1.01	GD
24400	-0.96	GD
24405	-0.11	NE
24410	-0.13	NE
24415	0.18	NE
24420	0.37	NE
24425	0.23	GA
24430	0.32	NE
24435	0.3	NE
24440	0.36	GA
24445	0.12	NE
24450	-0.3	NE
24455	0.11	NE
24460	0.2	NE
24465	0.2	NE
24470	0.11	NE
24475	0.11	NE
24480	0.37	GA
24485	0.99	NE
24490	0.95	GA
24495	0.89	GA
24500	0.61	GA
24505	-0.55	NE
24510	-inf	ES
24515	-inf	ES
24520	0.64	NE
24525	0.22	NE
24530	0.68	GA
24535	0.39	GA
24540	0.14	NE
24545	0.4	NE
24550	-inf	ES
24555	0.48	NE
24560	-inf	GD
24565	-inf	GD
24570	0.2	NE
24575	-0.34	NE
24580	-inf	ES
24585	0.12	NE
24595	-inf	GD
24600	-inf	ES
24605	0.11	NE
24610	0.25	NE
24615	-0.0	NE
24620	-0.14	NE
24625	-0.21	NE
24630	0.12	NE
24635	0.08	NE
24640	-0.83	GD
24645	-inf	ES
24650	0.15	NE
24655	0.34	NE
24660	0.43	NE
24665	0.55	NE
24670	0.22	NE
24675	0.35	NE
24680	-1.9	ES
24690	0.52	NE
24695	0.31	NE
24700	0.69	GA
24705	0.42	NE
24710	0.28	NE
24715	0.26	NE
24720	-0.2	NE
24725	0.05	NE
24730	-0.11	NE
24735	-0.15	NE
24740	0.06	NE
24745	0.11	NE
24750	0.33	NE
24755	-0.11	NE
24760	0.34	NE
24765	-0.02	NE
24770	0.47	NE
24775	-0.45	GD
24780	0.22	NE
24785	0.19	NE
24790	0.14	NE
24795	0.26	NE
24800	0.62	NE
24805	0.19	NE
24810	0.33	NE
24815	-0.08	NE
24820	-0.65	NE
24825	0.3	NE
24830	0.49	NE
24835	-0.42	NE
24840	0.69	NE
24845	0.31	NE
24850	0.28	NE
24855	0.05	NE
24860	0.34	NE
24865	0.5	GA
24870	0.36	NE
24875	-0.21	GD

Feature	Log(RF)	Class
24880	-0.12	NE
24885	0.3	NE
24890	-0.98	GD
24895	-0.45	GD
24900	-0.25	NE
24905	0.15	NE
24910	-0.17	NE
24915	-0.27	NE
24920	-0.58	GD
24925	-1.18	GD
24930	0.02	NE
24935	-0.04	NE
24940	0.29	NE
24945	-0.2	NE
24950	-0.38	GD
24955	-0.9	NE
24960	-0.49	GD
24965	-0.51	GD
24970	-1.26	GD
24975	0.03	NE
24980	0.82	GA
24985	0.51	NE
24990	0.41	NE
24995	0.39	NE
25000	0.15	NE
25005	0.25	NE
25015	0.2	NE
25020	-0.74	GD
25025	-0.2	NE
25030	-0.83	GD
25035	-0.21	GD
25040	-0.75	GD
25045	-0.12	NE
25050	0.04	NE
25055	0.07	NE
25060	0.14	NE
25065	0.47	NE
25070	0.23	NE
25075	0.46	NE
25080	0.13	NE
25085	0.74	NE
25090	-0.82	GD
25095	0.09	NE
25100	0.13	NE
25105	0.32	NE
25110	-0.08	NE
25115	0.07	NE
25120	0.02	NE
25125	-0.83	GD
25130	0.19	NE
25135	-0.32	GD
25140	-0.54	GD
25145	-0.2	GD
25150	0.12	NE
25155	-0.5	GD
25160	-0.74	GD
25165	0.6	NE
25170	0.09	NE
25175	0.28	NE
25180	-0.06	NE
25185	-0.3	NE
25190	-0.52	GD
25195	-0.64	GD
25200	0.1	NE
25205	-0.78	GD
25210	0.04	NE
25215	-0.58	GD
25220	0.1	NE
25225	-0.25	GD
25230	-0.07	NE
25235	0.32	NE
25240	0.34	GA
25245	0.37	NE
25250	0.81	GA
25255	0.45	NE
25260	0.5	GA
25265	-0.16	NE
25270	-0.42	GD
25275	0.03	NE
25280	0.04	NE
25285	0.19	NE
25290	0.48	NE
25295	0.24	NE
25300	-0.56	GD
25305	0.27	NE
25310	0.08	NE
25315	-0.01	NE
25325	-0.12	NE
25330	0.38	NE
25335	-0.03	NE
25340	-0.24	GD
25345	-0.29	GD
25350	-0.47	GD
25355	-0.44	GD
25360	0.05	NE
25365	-0.45	GD

Feature	Log(RF)	Class
25370	-0.18	GD
25375	-0.26	GD
25380	0.4	NE
25385	0.29	NE
25390	0.65	GA
25395	0.13	NE
25400	-0.02	NE
25405	0.21	NE
25410	-0.23	GD
25415	0.39	NE
25420	-0.03	NE
25425	-2.74	ES

Appendix B

Protocol for preparing Tn-seq libraries

Tn-seq Library Preparation-Q5

Version 14
Short Code: tnlibV14

Prepared by: Will Matern
Last updated: 04/01/2019 at 11:01

Equipment and Reagents

- TE (Tris EDTA) buffer
- Covaris MicroTube
- Covaris S220
- Tris-Cl buffer
- Nanodrop
- Thermo-Fisher Fast DNA End Repair Kit (Cat. #K0771)
- AxyPrep beads
- Magnetic plate (for removing beads)
- 70% EtOH (Ethanol)
- ThermoPol buffer
- dNTPs (NEB)
- Taq DNA Polymerase (NEB #M0267)
- 100uM adapter oligo 1 (ATGATGGCCGGTGGATTTGTGNNANNANNNTG-GTCGTGGTAT)
- 100uM adapter oligo 2 (pTACCACGACCA-NH_2, 5 prime phosphorylated with amino modifier at 3 prime)
- 50mM $MgCl_2$
- Thermocycler
- Blunt/TA Ligase Master Mix (NEB #M0367S)
- Q5 Hot Start 2X Master Mix
- 10uM adapter primer (ATGATGGCCGGTGGATTTGTG)
- 10uM transposon primer (TAATACGACTCACTATAGGGTCTAGAG)
- SPRIselect beads
- 85% EtOH
- 10uM sol_adapt [matches adapter, see below for sequence]
- 10uM sol_mar mix [matches transposon, see below for sequence]

Sequencing Oligos

10uM sol_mar mix consists of an equal parts mix of the following oligos diluted to 10uM in Tris-Cl:

- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGATCT CGGGGACTTATCAGCCAACC
- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGATCT TCGGGGACTTATCAGCCAACC
- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGATCT GATACGGGGACTTATCAGCCAACC
- AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGATCT ATCTACGGGGACTTATCAGCCAACC

10uM sol_adapt consits of *one* of the following oligos diluted to 10uM in Tris-Cl. Note that if multiplexing, a different oligo for each sample on the run must be used. Each oligo is constructed as one of four "backbones" with a different barcode (index) for distinguishing multiplexed samples. A mutliplexed run should consist of an equal molar mix of each backbone (to increase base diversity)

- Backbone 1: CAAGCAGAAGACGGCATACGAGAT XXX XXX XX GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTGTCAATGATGGCCGGTGGATT TGTG
- Backbone 2: CAAGCAGAAGACGGCATACGAGAT XXX XXX XX GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTCGTCCATGATGGCCGGTGGAT TTGTG
- Backbone 3: CAAGCAGAAGACGGCATACGAGAT XXX XXX XX GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTACAGTCCCATGATGGCCGGTG GATTTGTG
- Backbone 4: CAAGCAGAAGACGGCATACGAGAT XXX XXX XX GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTTAGTGGATGATGGCCGGTGGA TTTGTG

Where "XXX XXX XX" are barcodes used to distinguish mutliplexed samples. Barcodes we've used include the following: "ACA CGA TC", "AGC ATA CA", "TGC TAC GC", "AGT CTA CA", "CTC ATG CA", "AGT TCG GA", "CAT GAT CG", "CGT CAT CA", "CGC GCG GT", "GAC CTG CA", "TGA GAC TT", "AAG TAG AG", "GAG ATC TT", "GCC GAT GT", "TAC GTA CC", "CAG TTC AT", "TCC CTA TA", "GTC CGA TC", "GGT TCA AC", "CAC GTA CT", "TGT CAA GT", "TGT TCC GA", "TTC CGG AG", "CGA TCA AG", "CGA GGA GA", "TGG GGG AC", "TGC CTC GG", "TTA CAA CG", "CGA AAC CC", "ATC ACT CT", "TTC AGC AT", "ACT TGG TG". Note: The Illumina software to demultiplex will organize reads by the reverse complement of the sequence included in the above oligos.

Procedure

Shear DNA

1. Combine 5 ug (or more) of transposon containing genomic DNA with TE buffer to a total volume of 130uL in a Covaris MicroTube.
2. Settings for Covaris S220/E220 in NGSC for shearing to 500bp Note: These differ from settings suggested by Covaris. You should titrate the shear duration the first time and run a gel to confirm proper size. Duty Cycle-5%, Intensity-3, Cycle/Burst-200, Time-50 seconds. Degas water bath for at least 30 minutes. Bath temperature should reach ~5°C before shearing.
3. Purify DNA with spin column. **Elute into 43.5 uL of Tris-Cl**.
4. Remove 1uL to quantify DNA with nanodrop.

End-repair

1. Kit name: Thermo-Fisher Fast DNA End Repair Kit (Cat. #K0771)
2. Add the following to a clean PCR tube:

DNA	~3ug ( )
10X End Repair Buffer	5uL
DNA End Repair Enzyme Mix	2.5uL
Water	( )
Total Volume	50uL

3. Incubate at 20°C on lab bench for 5 minutes (do not exceed 20 min).
4. Purify DNA with **90uL** AxyPrep beads. **Elute into 32 uL of Tris-Cl** into a PCR tube.

AxyPrep Mag PCR Clean-up

1. Take AxyPrep Mag PCR Clean-up bead solution out of 4°C. Mix well to homogeneously disperse the beads before use.
2. Add 1.8x volume of beads to sample (18uL per 10uL of sample).
3. Mix beads and sample by pipetting 5 times.
4. Incubate mixed samples for **5 minutes** at room temperature.
5. Place samples onto magnetic plate for $\sim$ 1 minute, or until solution is clear.
6. Aspirate cleared solution from beads and discard.
7. Hint: Aspirate slowly from the very bottom of the tube and possibly leave a small amount behind to avoid pipetting the beads.
8. Wash beads with 180uL of 70% EtOH. Let sit for 30 seconds before aspirating ethanol and discarding. Repeat for a total of two washes.
9. Allow residual EtOH to evaporate for 1-5 minutes. Do not overdry!
10. Remove samples from magnetic plate. Add appropriate volume (**1-2uL more than you need**) of appropriate elution buffer. Let sit for 1 min to elute DNA.
11. Place samples back on magnetic plate to separate beads from solution.
12. Transfer eluate to fresh tube.

A-tailing

1. Using tube from A-tailing combine:

DNA	32uL
10X ThermoPol Buffer	5uL
10mM dATP	10uL
Taq DNA Polymerase (NEB: #M0267)	3uL
Total Volume	50uL

2. Incubate tube at 72.0 °C for 45 minutes in a thermocycler.
3. Purify with **90uL** AxyPrep beads. **Elute into 13uL of Tris-Cl**. Hint: You can preload the fresh tube with adapter (see: Ligation of adapters).

Adapter Annealing

1. Combine the following in a .2mL PCR tube:

100uM Adapter Oligo 1 (TACCA...)	48uL
100uM Adapter Oligo 2 (ATGATG...)	48uL
50mM $MgCl_2$	4uL
Total Volume	100uL

2. Using thermocycler, heat adapter mix to 95 °C for 10 minutes and then slowly reduce temperature to 20°C over 2 hours ($\sim$ .6 °C/min).

Ligation of adapters

1. To a fresh PCR tube, add the following (add master mix last and mix well):

annealed adapter mix	2uL
(A-tailed DNA)	13uL
Blunt/TA Ligase Master Mix (NEB #M0367S)	15uL
Total Volume	30uL

2. Incubate 60min at room temperature (on bench).
3. Purify with **54uL** AxyPrep beads. **Elute into 21uL of Tris-Cl**.
4. Quantify DNA with nanodrop. You should have .8-2 ug (40-100 ng/uL) of DNA.

PCR1: Himar1 + adapter

1. In a PCR tube combine the following. Hint: Add primers + mastermix to each PCR well first. Then add water and DNA to each well and pipette up-and-down to mix.

Reagent	×1
Q5 Hot Start 2X Master Mix	25uL
10uM adapter primer	2.5uL
10uM transposon primer	2.5uL
Adapter Ligated DNA	800ng ( )
Water	( )
Total volume per tube	50uL

2. Amplify DNA in thermocycler using the following PCR protocol:

1 cycle	30s @ 98 °C
20 cycles	10s @ 98 °C
	30s @ 65 °C
	30s @ 72 °C
1 cycle	2min @ 72 °C
	30 minutes

SPRIselect size selection (230bp-700bp)

1. Thoroughly shake SPRIselect bottle to resuspend beads. Add **27.5uL** of beads to 50uL PCR sample.
2. Mix total reaction volume by pipetting 10 times and incubate at room temperature for 1 minute.
3. Place tube on magnetic stand and allow beads to settle on magnet.
4. Transfer clear eluate to clean tube. Discard tubes with beads.
5. Thoroughly shake SPRIselect bottle to resuspend beads. Add additional **12.5uL** of beads to sample (total volume: 90uL).
6. Mix total reaction volume by pipetting 10 times and incubate at room temperature for 1 minute.
7. Place tube on magnetic stand and allow beads to settle on magnet.
8. Discard supernatant by pipetting from the bottom of the tube.
9. With tube still on magnet, add 180uL of 85% ethanol and incubate at room temperature for 30 seconds.
10. Discard ethanol supernatant by pipetting.
11. Allow residual ethanol to evaporate by leaving tube open on bench for 2 minutes.
12. Remove tube from magnet and add **50uL of Tris-Cl** to elute DNA. Mix by pipetting up and down 10 times and incubate at room temperature for 1 minute.
13. Place tube on magnetic stand and allow beads to settle.
14. Transfer clear eluate to eppendorf tube.
15. Measure samples with Nanodrop. Dilute all samples to between 0.4 - 1.25 ng/uL. Note: Preliminary evidence suggests that Qubit greatly underestimates the amount of DNA at this step (5-10 fold), possibly due to large amounts of ssDNA. Note: Dilution is needed in order to avoid saturating the PCR#2 reaction (which causes PCR bubbles - making BioA useless).

PCR2: Hemi-nested PCR and Addition of Illumina attachment sequences (P5/P7)

1. In a PCR tube combine the following:

Reagent	×1
10uM sol_adapt (matches adapter)	1uL
10uM sol_mar mix (matches Tn)	1uL
Diluted size-selected DNA (0.4 - 1.25 ng/uL)	2uL
Q5 Hot Start 2X Master Mix	10uL
Water	6uL
Total volume per tube	20uL

2. Amplify DNA in thermocycler using the following protocol:

1 cycle	30s @ 98 °C
10 cycles	10s @ 98 °C
	30s @ 67 °C
	30s @ 72 °C
1 cycle	2min @ 72 °C
	15 minutes

3. Purify with **18uL** SPRIselect beads. **Elute into 20uL Tris-Cl**.
4. Quantify DNA with Qubit or qPCR (Nanodrop is NOT reliable). DNA concentration should be between 5 - 20 nM (and must be at least 2nM for sequencing). If concentration is not high enough then increase input to PCR#2. Note: On Agilent Bioanalyzer traces these libraries gave a small peak at about twice the size of the rest of the library (1000bp vs 500bp). This is believed to be due to some non-specific amplicon (likely linear amplification from the adapter sequence). This secondary peak should not affect bulk quantitation via Qubit, Bioanalyzer, or qPCR. Hypothetical ways to eliminate these peaks include increasing cycles of PCR#2. Note: Qubit is faster and easier for quantitation.

Appendix C

Hypersusceptibility predictions for Mav

The tables on the following pages report predictions for the impact of disrupting each of the annotated genomic features in Mav strain MAC109 based on my collected Tn-seq data. Only statistically significant genes are reported (see Ch. 3 for thresholds). The four tables report the effect size of removing the gene on bacterial viability after exposure to CLR (5.4ug/mL), MOX (10ug/mL), RFB (0.63ug/mL), and EMB (2.1ug/mL) respectively. Negative values indicate the mutant is hypersusceptible to the drug while positive values indicate hypertolerance. Each feature is labelled with the final 5 digits of the locus tag provided in the Genbank file (see Ch.2 for reference). This number is unique for each feature and only excludes an invariant alphanumeric string to distinguish the genomes from other in Genbank (ie it excludes "DFS55_" specific to the MAC109 genome). The second and third columns provides the logarithm, base 2, of the fold-change of the mutant 12 and 48 hours of drug exposure, respectively (relative to no drug control). See chapter 3 for additional details. Computer-readable versions of these tables including additional information are provided as associated files (AppendixC_FullTable_CLR.csv, AppendixC_FullTable_MOX.csv, AppendixC_FullTable_RFB.csv, AppendixC_FullTable_EMB.csv).

Feature	CLR 12h	CLR 48h
03215	-2.88	-2.99
10120	-1.17	-1.91
12730	-1.57	-1.65
08190	-1.57	-1.39
15065	-0.9	-1.39
10815	-1.82	-1.32
16070	-1.16	-1.23
00905	-0.64	-1.22
03080	-0.82	-1.22
19360	-0.84	-1.19
02655	-1.31	-1.17
25370	-0.68	-1.1
07105	-0.71	-1.06
18590	-0.82	-1.04
08650	-1.01	-1.03
12665	-0.58	-1.02
08575	-1.62	-1.01
02650	-1.85	-0.99
02660	-1.76	-0.99
14910	-0.68	-0.9
18980	-1.51	-0.89
07230	-0.73	-0.89
05615	-0.92	-0.88
05685	-1.06	-0.87
00975	-1.79	-0.86
02265	-0.61	-0.86
23995	-0.54	-0.85
18480	-1.36	-0.84
14565	-0.67	-0.81
20365	-1.0	-0.78
25405	-0.97	-0.73
07130	-0.67	-0.72
25420	-0.95	-0.72
02440	-0.66	-0.72
25375	-0.63	-0.72
22340	-0.75	-0.72
25385	-1.01	-0.71
25380	-1.04	-0.71
17715	-1.03	-0.7
14655	-0.51	-0.68
21330	-0.52	-0.6
22280	-0.58	-0.59
23935	-0.58	-0.59
20105	-0.57	-0.57
24640	-0.62	-0.55
00980	-0.94	-0.54
25400	-0.6	-0.53
02385	-0.58	-0.52
13810	-0.55	-0.51
15050	-0.91	-0.51
16845	0.72	0.66
20235	1.08	0.73
18380	0.55	0.86
22165	0.54	0.93

Feature	CLR 12h	CLR 48h
05640	0.58	1.0
23890	0.63	1.01
10665	1.01	1.09
05665	0.88	1.1
20040	0.64	1.1
21755	0.52	1.28
01380	0.79	1.31
03015	0.52	1.39
21750	0.62	1.4
10765	1.18	1.48
13800	0.8	1.65
10660	1.5	1.91

Feature	MOX 12h	MOX 48h
23975	-2.67	-4.56
07350	-2.79	-4.47
10120	-3.1	-4.3
03885	-4.35	-3.65
15265	-2.05	-3.43
22595	-1.09	-3.18
14115	-1.42	-3.03
14295	-1.84	-2.97
07355	-1.61	-2.92
18595	-1.5	-2.73
10115	-1.61	-2.71
04840	-1.94	-2.69
22980	-1.93	-2.51
07940	-1.55	-2.5
23390	-1.56	-2.46
14300	-1.78	-2.43
08050	-1.36	-2.4
05660	-1.64	-2.4
05655	-1.65	-2.38
08530	-1.7	-2.31
20215	-1.47	-2.23
15065	-1.48	-2.21
21405	-0.57	-2.19
00905	-1.6	-2.14
20220	-1.67	-2.08
12730	-1.44	-1.92
19985	-1.17	-1.91
19835	-1.3	-1.9
21345	-1.6	-1.9
17245	-0.92	-1.9
19360	-1.17	-1.84
12835	-0.71	-1.83
00125	-0.97	-1.8
05695	-0.86	-1.78
22990	-0.62	-1.77
14275	-0.85	-1.73
09485	-1.72	-1.73
05375	-0.73	-1.71
07230	-1.16	-1.71
00355	-0.92	-1.69
00970	-1.69	-1.69
00015	-1.56	-1.68
05330	-0.94	-1.68
07105	-1.07	-1.66
12540	-0.88	-1.65
20210	-1.4	-1.64
07130	-1.09	-1.62
12530	-0.7	-1.62
18280	-1.1	-1.6
05720	-1.5	-1.59
07240	-0.94	-1.59
10110	-1.15	-1.57
00900	-0.99	-1.55
14810	-0.94	-1.54
14355	-0.7	-1.53
05715	-0.54	-1.52
13140	-0.81	-1.49
11275	-0.67	-1.48
16635	-1.0	-1.47
18275	-0.81	-1.45
12535	-0.84	-1.43
07235	-0.61	-1.43
22485	-0.5	-1.43
01570	-0.92	-1.42
10925	-0.98	-1.42
23895	-0.92	-1.4
04845	-0.96	-1.39
10470	-0.63	-1.37
05325	-0.6	-1.34
25390	-0.8	-1.32
23495	-0.79	-1.32
05685	-0.95	-1.31
03500	-0.64	-1.29
07785	-0.65	-1.25
18915	-0.74	-1.22
02950	-1.17	-1.22
16630	-1.23	-1.22
03200	-0.53	-1.22
02640	-0.78	-1.2
12510	-0.79	-1.17
14865	-0.69	-1.16
25370	-0.72	-1.13
12525	-0.58	-1.13
10400	-0.78	-1.12
10405	-0.69	-1.1
03125	-0.54	-1.08
00360	-0.72	-1.07
23740	-0.88	-1.06
12665	-0.74	-1.01
19670	-0.64	-1.01
02800	-0.64	-0.97
19515	-0.54	-0.97
19850	-0.52	-0.95
10575	-0.75	-0.95
04450	-0.53	-0.94
03530	-0.9	-0.94
20355	-0.62	-0.92
15050	-0.56	-0.89
02440	-0.6	-0.8
14095	-0.91	-0.78
14530	-0.51	-0.77
19520	-0.62	-0.69
18590	-0.7	-0.68
19220	-0.7	-0.67
14915	-0.63	-0.56
23935	-0.57	-0.53
05640	0.51	0.56
20040	0.75	1.32
10765	0.74	1.48

Feature	RFB 12h	RFB 48h
03215	-2.75	-3.68
25390	-1.56	-2.47
15065	-1.43	-2.17
10120	-2.6	-2.03
25370	-1.1	-1.93
25360	-1.98	-1.89
10125	-2.36	-1.87
20215	-0.9	-1.87
15265	-1.64	-1.73
03080	-1.52	-1.72
11285	-0.95	-1.7
21345	-1.02	-1.69
18595	-1.22	-1.69
20220	-1.23	-1.64
19360	-1.11	-1.63
07350	-2.35	-1.62
23975	-2.92	-1.58
05660	-0.97	-1.55
12540	-0.5	-1.51
10470	-0.54	-1.47
12730	-1.55	-1.44
25410	-1.18	-1.37
22480	-0.84	-1.37
08215	-0.77	-1.34
00970	-1.19	-1.32
03200	-0.63	-1.3
25420	-1.06	-1.28
07130	-0.73	-1.19
24570	-0.83	-1.17
25385	-0.88	-1.17
05330	-0.73	-1.17
25405	-0.98	-1.16
25375	-0.73	-1.16
14660	-1.41	-1.15
00125	-0.57	-1.15
00355	-1.96	-1.09
14810	-0.77	-1.09
00360	-0.94	-1.08
10115	-1.89	-1.07
07355	-1.99	-1.06
05695	-0.95	-1.01
20210	-1.26	-1.0
07105	-0.58	-0.99
25365	-0.86	-0.98
18590	-1.78	-0.96
08650	-2.25	-0.93
23390	-1.3	-0.92
00900	-0.75	-0.9
08725	-0.5	-0.9
13390	-0.55	-0.89
07240	-0.88	-0.89
08530	-0.51	-0.88
10925	-1.52	-0.86
20500	-0.97	-0.82

Feature	RFB 12h	RFB 48h
14115	-1.82	-0.79
00905	-1.11	-0.79
19835	-0.86	-0.79
25395	-0.64	-0.79
25415	-0.94	-0.78
11595	-0.58	-0.77
02385	-0.63	-0.77
05655	-0.87	-0.75
22990	-0.76	-0.74
25400	-0.69	-0.74
02905	-0.58	-0.72
11005	-0.67	-0.71
00120	-0.84	-0.7
14275	-0.81	-0.68
15050	-0.58	-0.68
18840	-1.05	-0.67
02605	-0.59	-0.64
05190	-0.64	-0.61
05685	-1.46	-0.6
04985	-0.71	-0.6
19670	-0.69	-0.6
03055	-0.5	-0.59
20495	-1.09	-0.57
19940	-0.62	-0.56
07230	-0.89	-0.54
03890	-0.67	-0.54
02800	-0.72	-0.53
10575	-0.93	-0.52
08665	0.78	0.52
19605	0.52	0.58
02500	0.54	0.73
10205	0.87	0.81
10665	0.94	0.83
18875	0.71	0.84
01580	0.63	0.88
21750	0.69	0.91
14525	0.58	0.94
19735	0.89	0.99
10325	0.57	1.09
02115	0.61	1.21
21365	0.75	1.23
07565	0.9	1.34
20040	1.18	1.4
05640	1.44	1.41
12705	0.65	1.43
05645	1.19	1.55
03545	1.19	1.76
10660	1.41	1.93
16845	1.39	2.19
10765	2.01	2.47

Feature	EMB 12h	EMB 48h
00905	-0.72	-1.4
10120	-0.54	-1.3
03885	-0.68	-1.2
00120	-0.87	-1.19
09485	-0.98	-1.14
19515	-0.83	-1.1
12730	-0.56	-0.83
00360	-0.59	-0.82
08650	-0.54	-0.72

Appendix D

Hypersusceptibility predictions for Mtb

The tables on the following pages report predictions for the impact of disrupting each of the annotated genomic features in Mtb strain H37Rv based on collected Tn-seq data. Only statistically significant genes are reported (see Ch. 4 for thresholds). The four tables report the effect size of removing the gene on bacterial viability after exposure to INH (1.0ug/mL) or RMP (4.0ug/mL). Tables for the rich medium condition are printed first followed by tables for the starvation condition. Negative values indicate the mutant is hypersusceptible to the drug while positive values indicate hypertolerance. The first column provides the name of each feature. The second and subsequent columns provides the logarithm, base 2, of the fold-change of the mutant after 6d (for rich medium) or 7d and 14d (for starvation) relative to no drug control. See chapter 4 for additional details. Computer-readable versions of these tables including additional information are provided as associated files (AppendixD_FullTable_7H9_INH.csv, AppendixD_FullTable_7H9_RMP.csv, AppendixC_FullTable_PBS_INH.csv, AppendixC_FullTable_PBS_RMP.csv).

Feature	INH
Rv2428	-1.83
Rv1901	-1.22

Feature	RMP
Rv0994	-3.25
Rv0049	-2.99
Rv2179c	-2.72
Rv0472c	-2.29
Rv1433	-2.08
Rv2047c	-1.87
Rv2224c	-1.80
Rv3267	-1.79
Rv2525c	-1.78
Rv2190c	-1.69
Rv0862c	-1.68
Rv0999	-1.65
Rv1411c	-1.65
Rv0503c	-1.64
Rv2781c	-1.58
Rv1244	-1.56
Rv1821	-1.54
Rv3207c	-1.52
Rv1860	-1.50
Rv1435c	-1.49
Rv0202c	-1.49
Rv1127c	-1.46
Rv3717	-1.46
Rv0861c	-1.45
Rv1836c	-1.45
Rv1126c	-1.44
Rv0820	-1.42
Rv3822	-1.39
Rv2363	-1.35
Rv3811	-1.32
Rv1006	-1.28
Rv0497	-1.27
Rv3005c	-1.25
Rv2700	-1.25
Rv1410c	-1.23
Rv0929	-1.23
Rv1884c	-1.22
Rv1566c	-1.22
Rv2171	-1.16
Rv0930	-1.16
Rv1886c	-1.13
Rv2176	-1.11
Rv1272c	-1.07
Rv0129c	-1.05
Rv1273c	-1.05
Rv1683	-1.05
Rv0642c	-1.04
Rv1096	-1.03
Rv3330	-0.99
Rv0309	-0.99
Rv3922c	-0.94
Rv3682	-0.94
Rv0474	-0.92
Rv0928	-0.90
Rv3630	-0.87
Rv3919c	-0.87
Rv1608c	-0.86
Rv3311	-0.86
Rv2248	-0.84
Rv3818	-0.83
Rv1209	-0.82
Rv1421	-0.82
Rv2463	-0.82
Rv1282c	-0.82
Rv3719	-0.81
Rv2696c	-0.80
Rv2048c	-0.80
Rv0007	-0.79
Rv2170	-0.77
Rv1280c	-0.77
Rv1281c	-0.72
Rv1607	-0.72
Rv1555	-0.72
Rv3803c	-0.71
Rv2672	-0.71
Rv2197c	-0.71
Rv0179c	-0.69
Rv2553c	-0.68
Rv0805	-0.67
Rv3178a	-0.66
Rv1246c	-0.66
Rv3600c	-0.66
Rv0111	-0.65
Rv0500A	-0.65
Rv1159	-0.65
Rv2684	-0.65
Rv3632	-0.65
Rv3271c	-0.64
Rv3910	-0.64
Rv0961	-0.64
Rv2931	-0.63
Rv2508c	-0.62
Rv0996	-0.61
Rv3779	-0.61
Rv1459c	-0.61
Rv2462c	-0.59
Rv1220c	-0.58
Rv2398c	-0.58
Rv0051	-0.58
Rv0204c	-0.58
Rv0017c	-0.57
Rv1698	-0.57
Rv2930	-0.56
Rv0954	-0.56
Rv2214c	-0.56
Rv0191	-0.55
Rv3775	-0.55
Rv0432	-0.55
Rv0201c	-0.55
Rv1420	-0.54
Rv2933	-0.54
Rv2864c	-0.54
Rv1478	-0.54
Rv3256c	-0.53
Rv0235c	-0.53
Rv3820c	-0.53
Rv3310	-0.53
Rv1541c	-0.53
Rv0009	-0.52
Rv0483	-0.51
Rv1057	-0.50
Rv3869	0.50
Rv0392c	0.51
Rv3057c	0.54
Rv3873	0.56
Rv1070c	0.57
Rv0092	0.60
Rv1333	0.60
Rv1197	0.61
Rv2334	0.61
Rv0362	0.61
Rv1334	0.64
Rv0812	0.65
Rv2372c	0.65
Rv3316	0.67
Rv0167	0.67
Rv0530A	0.69
Rv2404c	0.69
Rv3876	0.72
Rv1337	0.72
Rv0056	0.74
Rv2772c	0.74
Rv2970c	0.74
Rv3615c	0.74
Rv2347c	0.75
Rv0157	0.75
Rv0178	0.76
Rv3777	0.77
Rv1486c	0.77
Rv1956	0.78
Rv0238	0.79
Rv2222c	0.80
Rv3871	0.82
Rv3269	0.83
Rv3135	0.85
Rv2722	0.85
Rv2335	0.85
Rv1908c	0.85
Rv1008	0.86
Rv2326c	0.88
Rv1129c	0.90
Rv0169	0.91
Rv2702	0.91
Rv0175	0.94
Rv3842c	0.94
Rv1009	0.94
Rv1335	0.94
Rv0905	0.94
Rv2392	0.95
Rv0734	0.95
Rv1745c	0.95
Rv0171	0.99
Rv1030	0.99
Rv3421c	1.00
Rv0513	1.01
Rv1029	1.01
Rv2173	1.02
Rv0172	1.03
Rv0177	1.03
Rv3270	1.05
Rv0818	1.10
Rv2782c	1.10
Rv1130	1.10
Rv0199	1.14
Rv2721c	1.15
Rv0168	1.15
Rv0174	1.15
Rv1336	1.16
Rv2709	1.16
Rv0173	1.17
Rv3723	1.18
Rv1455	1.18
Rv0508	1.21
Rv1013	1.24
Rv0176	1.24
Rv0170	1.25
Rv1112	1.31
Rv1691	1.32
Rv1170	1.32
Rv0655	1.32
Rv0819	1.33
Rv2391	1.49
Rv1387	1.57
Rv3419c	1.57
Rv0200	1.61
Rv3484	1.69
Rv1386	2.01
Rv2140c	2.13
Rv2691	2.15
Rv2692	2.18
Rv2694c	2.18
Rv1957	2.26
Rv3200c	2.84
Rv2690c	5.56

Feature	INH 7d	INH 14d
Rv1901	-3.15	-3.22
Rv0767c	-2.29	-2

Feature	RMP 7d	RMP 14d	Feature	RMP 7d	RMP 14d
Rv0049	-1.45	-1.81	Rv0694	1.23	0.77
Rv2374c	-1.47	-1.68	Rv0169	1.96	0.77
Rv0819	-1.44	-1.67	Rv0171	1.98	0.77
Rv3680	-1.12	-1.40	Rv0177	2.10	0.78
Rv3923c	-1.10	-1.36	Rv0174	2.26	0.78
Rv2179c	-1.32	-1.30	Rv0176	2.06	0.78
Rv0998	-1.24	-1.25	Rv0170	2.05	0.80
Rv2051c	-1.33	-1.23	Rv0693	1.66	0.81
Rv2733c	-1.55	-1.19	Rv3311	1.41	0.85
Rv2392	-1.29	-1.17	Rv0692	1.87	0.86
Rv2709	-1.26	-1.16	Rv0513	3.12	0.92
Rv1908c	-0.88	-1.15	Rv0863	0.70	0.94
Rv0994	-1.27	-1.01	Rv3331	0.61	0.94
Rv0956	-1.11	-0.95	Rv0178	2.27	0.95
Rv3624c	-0.96	-0.92	Rv2173	0.73	0.98
Rv0238	-0.65	-0.90	Rv0989c	2.77	1.01
Rv3261	-0.88	-0.87	Rv2633c	1.76	1.02
Rv1836c	-1.14	-0.84	Rv0458	2.30	1.05
Rv1328	-0.57	-0.82	Rv1183	0.52	1.12
Rv2731	-1.00	-0.81	Rv0678	1.24	1.43
Rv1638	-0.74	-0.79	Rv2690c	3.26	1.63
Rv1691	-0.75	-0.78	Rv2199c	2.53	2.42
Rv0003	-0.52	-0.74	Rv0199	4.50	2.53
Rv0949	-1.07	-0.74	Rv0200	4.69	2.56
Rv3024c	-1.01	-0.72	Rv0655	5.03	3.01
Rv3160c	-0.80	-0.70	Rv3723	4.94	3.04
Rv1624c	-0.53	-0.69			
Rv1633	-0.59	-0.68			
Rv2793c	-0.80	-0.66			
Rv2737c	-0.80	-0.65			
Rv1909c	-0.58	-0.64			
Rv2391	-0.72	-0.63			
Rv0861c	-0.76	-0.63			
Rv0999	-0.82	-0.60			
Rv2672	-1.01	-0.59			
Rv0680c	-0.57	-0.56			
Rv3544c	-0.74	-0.55			
Rv0530A	-0.74	-0.55			
Rv3241c	-0.57	-0.55			
Rv2047c	-1.16	-0.54			
Rv2048c	-0.92	-0.50			
Rv0503c	0.54	0.52			
Rv0761c	1.46	0.52			
Rv2536	1.73	0.57			
Rv1016c	1.82	0.58			
Rv0696	1.49	0.59			
Rv2500c	0.76	0.64			
Rv0175	2.15	0.64			
Rv0172	2.11	0.73			
Rv1211	1.69	0.73			
Rv0173	2.20	0.75			
Rv1017c	2.05	0.75			
Rv0167	1.87	0.76			
Rv0168	1.94	0.77			

Appendix E

Essential genes shared between Mav and Mtb

The table on the next page reports genes found to be essential in both Mav and Mtb. A computer-readable version of this table including additional information is provided as an associated file (AppendixE_FullTable.csv).

MAC109 Name	H37Rv Name
DFS55_00005	Rv0001
DFS55_00025	Rv0005
DFS55_00030	Rv0006
DFS55_00095	Rv0014c
DFS55_00280	Rv0041
DFS55_00340	Rv0046c
DFS55_00440	Rv3907c
DFS55_01330	Rv0164
DFS55_01590	Rv0206c
DFS55_01700	Rv0224c
DFS55_01705	Rv0225
DFS55_01715	Rv0227c
DFS55_01805	Rv0236c
DFS55_02315	Rv0334
DFS55_02330	Rv0338c
DFS55_02375	Rv0350
DFS55_02490	Rv0364
DFS55_02525	Rv0357c
DFS55_02675	Rv0415
DFS55_02760	Rv0423c
DFS55_02880	Rv0440
DFS55_03250	Rv0509
DFS55_03265	Rv0512
DFS55_03345	Rv0524
DFS55_03835	Rv0635
DFS55_03840	Rv0636
DFS55_03860	Rv0639
DFS55_03955	Rv0667
DFS55_03960	Rv0668
DFS55_04035	Rv0684
DFS55_04040	Rv0685
DFS55_04125	Rv0700
DFS55_04130	Rv0701
DFS55_04135	Rv0702
DFS55_04165	Rv0708
DFS55_04230	Rv0718
DFS55_04235	Rv0719
DFS55_04320	Rv0732
DFS55_04460	Rv3464
DFS55_04475	Rv3462c
DFS55_04485	Rv3460c
DFS55_04495	Rv3458c
DFS55_04560	Rv3443c
DFS55_04570	Rv3441c
DFS55_04605	Rv3436c
DFS55_04630	Rv3423c
DFS55_04705	Rv3411c
DFS55_04745	Rv3396c
DFS55_04920	Rv3336c
DFS55_05180	Rv3285
DFS55_05205	Rv3280
DFS55_05300	Rv3266c
DFS55_05305	Rv3265c
DFS55_05310	Rv3264c
DFS55_05355	Rv3257c
DFS55_05365	Rv3255c
DFS55_05400	Rv3248c
DFS55_05405	Rv3247c
DFS55_05435	Rv3240c
DFS55_05690	Rv3206c
DFS55_05740	Rv3198c
DFS55_06620	Rv3105c
DFS55_06645	Rv3100c
DFS55_07025	Rv3053c
DFS55_07030	Rv3052c
DFS55_07035	Rv3051c
DFS55_07060	Rv3048c
DFS55_07080	Rv3043c
DFS55_07285	Rv3014c
DFS55_07330	Rv3011c
DFS55_07345	Rv3009c
DFS55_07410	Rv2992c
DFS55_07470	Rv2981c
DFS55_07555	Rv2969c
DFS55_08005	Rv2883c
DFS55_08045	Rv2870c
DFS55_08180	Rv2845c
DFS55_08210	Rv2839c
DFS55_08285	Rv2786c
DFS55_08295	Rv2783c
DFS55_08415	Rv2764c
DFS55_08470	Rv2748c
DFS55_08690	Rv2703
DFS55_08775	Rv2682c
DFS55_08795	Rv2678c
DFS55_08800	Rv2677c
DFS55_08805	Rv2676c
DFS55_08825	Rv2673
DFS55_09180	Rv2614c
DFS55_09195	Rv2611c
DFS55_09330	Rv2580c
DFS55_09395	Rv2572c
DFS55_09475	Rv2555c
DFS55_09540	Rv2540c
DFS55_09545	Rv2539c
DFS55_09550	Rv2538c
DFS55_09575	Rv2533c
DFS55_09675	Rv1383
DFS55_09680	Rv1384
DFS55_09715	Rv1392
DFS55_09750	Rv1402
DFS55_09795	Rv1412
DFS55_09800	Rv1415
DFS55_09980	Rv1449c
DFS55_10025	Rv1461
DFS55_10030	Rv1462
DFS55_10035	Rv1463
DFS55_10100	Rv1475c
DFS55_10145	Rv1484
DFS55_10150	Rv1485
DFS55_10445	Rv1536
DFS55_10500	Rv1547
DFS55_10730	Rv1609
DFS55_10745	Rv1612
DFS55_10750	Rv1613
DFS55_10755	Rv1614
DFS55_10845	Rv1630
DFS55_11030	Rv1650
DFS55_11035	Rv1652
DFS55_11045	Rv1654
DFS55_11050	Rv1655
DFS55_11065	Rv1658
DFS55_11070	Rv1659
DFS55_11215	Rv1689
DFS55_11245	Rv1695
DFS55_11265	Rv1699
DFS55_22305	Rv3596c
DFS55_12735	Rv1837c
DFS55_12795	Rv1854c
DFS55_14110	Rv2051c
DFS55_14370	Rv2121c
DFS55_14550	Rv2139
DFS55_14665	Rv2145c
DFS55_14675	Rv2147c
DFS55_14705	Rv2152c
DFS55_14710	Rv2153c
DFS55_14735	Rv2158c
DFS55_14740	Rv2163c
DFS55_14750	Rv2165c
DFS55_14785	Rv2174
DFS55_14825	Rv2182c
DFS55_14880	Rv2192c
DFS55_14885	Rv2193
DFS55_14895	Rv2195
DFS55_14900	Rv2196
DFS55_14930	Rv2201
DFS55_15035	Rv2220
DFS55_15210	Rv2242
DFS55_15225	Rv2245
DFS55_16075	Rv2343c
DFS55_16105	Rv2357c
DFS55_16125	Rv2361c
DFS55_21315	Rv0824c
DFS55_16660	Rv2421c
DFS55_16725	Rv2438c
DFS55_16750	Rv2439c
DFS55_16785	Rv2447c
DFS55_16790	Rv2448c
DFS55_16825	Rv2457c
DFS55_16935	Rv2477c
DFS55_17230	Rv2524c
DFS55_17765	Rv1327c

MAC109 Name	H37Rv Name
DFS55_17770	Rv1326c
DFS55_17900	Rv1310
DFS55_17905	Rv1309
DFS55_17910	Rv1308
DFS55_17915	Rv1307
DFS55_17940	Rv1302
DFS55_17955	Rv1299
DFS55_17965	Rv1297
DFS55_17975	Rv1296
DFS55_17985	Rv1294
DFS55_17990	Rv1293
DFS55_17995	Rv1292
DFS55_18395	Rv1254
DFS55_18675	Rv1202
DFS55_18690	Rv1201c
DFS55_18870	Rv1166
DFS55_19275	Rv1098c
DFS55_19300	Rv1094
DFS55_19305	Rv1093
DFS55_19310	Rv1092c
DFS55_19580	Rv1023
DFS55_19620	Rv1017c
DFS55_19680	Rv1007c
DFS55_19995	Rv0982
DFS55_20065	Rv0957
DFS55_20385	Rv0884c
DFS55_21410	Rv0811c
DFS55_21420	Rv0809
DFS55_21425	Rv0808
DFS55_21440	Rv0803
DFS55_21540	Rv0780
DFS55_21555	Rv0777
DFS55_22220	Rv3580c
DFS55_22235	Rv3583c
DFS55_22255	Rv3587c
DFS55_22315	Rv3598c
DFS55_22330	Rv3602c
DFS55_22440	Rv3628
DFS55_22465	Rv3634c
DFS55_22470	Rv3635
DFS55_22490	Rv3646c
DFS55_22950	Rv3709c
DFS55_22970	Rv3712
DFS55_22975	Rv3713
DFS55_23010	Rv3721c
DFS55_23015	Rv3722c
DFS55_23710	Rv3781
DFS55_23715	Rv3782
DFS55_23745	Rv3790
DFS55_23750	Rv3791
DFS55_23755	Rv3792
DFS55_23760	Rv3793
DFS55_23780	Rv3794
DFS55_23785	Rv3795

MAC109 Name	H37Rv Name
DFS55_23810	Rv3799c
DFS55_23815	Rv3800c
DFS55_23820	Rv3801c
DFS55_23840	Rv3805c
DFS55_23845	Rv3806c
DFS55_23855	Rv3808c
DFS55_23860	Rv3809c
DFS55_23930	Rv3834c
DFS55_24050	Rv3858c
DFS55_24055	Rv3859c
DFS55_24510	Rv3909
DFS55_24515	Rv3910
DFS55_24550	Rv3915
DFS55_24580	Rv3921c

References

[1] Jan Abendroth et al. "Mycobacterium Tuberculosis Rv2179c Protein Establishes a New Exoribonuclease Family with Broad Phylogenetic Distribution". In: *The Journal of Biological Chemistry* 289.4 (24, 2014), pp. 2139–2147. ISSN: 1083-351X. DOI: 10.1074/jbc.M113.525683. pmid: 24311791.

[2] Katherine A. Abrahams et al. "Inhibiting Mycobacterial Tryptophan Synthase by Targeting the Inter-Subunit Interface". In: *Scientific Reports* 7.1 (25, 2017), p. 9430. ISSN: 2045-2322. DOI: 10.1038/s41598-017-09642-y. pmid: 28842600.

[3] Rebecca Joy Archuleta, Patricia Yvonne Hoppes, and Todd P. Primm. "Mycobacterium Avium Enters a State of Metabolic Dormancy in Response to Starvation". In: *Tuberculosis (Edinburgh, Scotland)* 85.3 (2005), pp. 147–158. ISSN: 1472-9792. DOI: 10.1016/j.tube.2004.09.002. pmid: 15850753.

[4] Rebecca Joy Archuleta, Patricia Yvonne Hoppes, and Todd P. Primm. "Mycobacterium Avium Enters a State of Metabolic Dormancy in Response to Starvation". In: *Tuberculosis (Edinburgh, Scotland)* 85.3 (2005), pp. 147–158. ISSN: 1472-9792. DOI: 10.1016/j.tube.2004.09.002. pmid: 15850753.

[5] N. Q. Balaban. "Bacterial Persistence as a Phenotypic Switch". In: *Science* 305.5690 (10, 2004), pp. 1622–1625. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1099390. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.1099390 (visited on 03/28/2019).

[6] Anton Bankevich et al. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing". In: *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 19.5 (2012), pp. 455–477. ISSN: 1557-8666. DOI: 10.1089/cmb.2012.0021. pmid: 22506599.

[7] J. T. Belisle and M. G. Sonnenberg. "Isolation of Genomic DNA from Mycobacteria". In: *Methods in Molecular Biology (Clifton, N.J.)* 101 (1998), pp. 31–44. ISSN: 1064-3745. DOI: 10.1385/0-89603-471-2:31. pmid: 9921467.

[8] Debra Benator et al. "Rifapentine and Isoniazid Once a Week versus Rifampicin and Isoniazid Twice a Week for Treatment of Drug-Susceptible Pulmonary Tuberculosis in HIV-Negative Patients: A Randomised Clinical Trial". In: *Lancet (London, England)* 360.9332 (17, 2002), pp. 528–534. ISSN: 0140-6736. pmid: 12241657.

[9] Yoav Benjamini and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1 (1995), pp. 289–300. ISSN: 00359246. JSTOR: 2346101.

[10] T. F. Blaschke and M. H. Skinner. "The Clinical Pharmacokinetics of Rifabutin". In: *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 22 Suppl 1 (1996), S15–21; discussion S21–22. ISSN: 1058-4838. pmid: 8785251.

[11] Anthony M. Bolger, Marc Lohse, and Bjoern Usadel. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data". In: *Bioinformatics (Oxford, England)* 30.15 (1, 2014), pp. 2114–2120. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btu170. pmid: 24695404.

[12] Laure Botella et al. “Depleting Mycobacterium Tuberculosis of the Transcription Termination Factor Rho Causes Pervasive Transcription and Rapid Death”. In: *Nature Communications* 8 (28, 2017), p. 14731. ISSN: 2041-1723. DOI: 10.1038/ncomms14731. pmid: 28348398.

[13] Lawrence Broxmeyer et al. “Killing of Mycobacterium Avium and Mycobacterium Tuberculosis by a Mycobacteriophage Delivered by a Nonvirulent Mycobacterium: A Model for Phage Therapy of Intracellular Bacterial Pathogens”. In: *The Journal of Infectious Diseases* 186.8 (15, 2002), pp. 1155–1160. ISSN: 0022-1899. DOI: 10.1086/343812. pmid: 12355367.

[14] Heike Brötz-Oesterhelt et al. “Specific and Potent Inhibition of NAD+-Dependent DNA Ligase by Pyridochromanones”. In: *The Journal of Biological Chemistry* 278.41 (10, 2003), pp. 39435–39442. ISSN: 0021-9258. DOI: 10.1074/jbc.M306479200. pmid: 12867414.

[15] Talha Khan Burki. “The Global Cost of Tuberculosis”. In: *The Lancet Respiratory Medicine* 6.1 (2018), p. 13. ISSN: 22132600. DOI: 10.1016/S2213-2600(17)30468-X. URL: https://linkinghub.elsevier.com/retrieve/pii/S221326001730468X (visited on 04/03/2019).

[16] E. A. Campbell et al. “Structural Mechanism for Rifampicin Inhibition of Bacterial Rna Polymerase”. In: *Cell* 104.6 (23, 2001), pp. 901–912. ISSN: 0092-8674. pmid: 11290327.

[17] George Carter et al. “Characterization of Biofilm Formation by Clinical Isolates of Mycobacterium Avium”. In: *Journal of Medical Microbiology* 52 (Pt 9 2003), pp. 747–752. ISSN: 0022-2615. DOI: 10.1099/jmm.0.05224-0. pmid: 12909649.

[18] P. Maureen Cassidy et al. “Nontuberculous Mycobacterial Disease Prevalence and Risk Factors: A Changing Epidemiology”. In: *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 49.12 (15, 2009), e124–129. ISSN: 1537-6591. DOI: 10.1086/648443. pmid: 19911942.

[19] Nancy Cheng et al. “Filtration Improves the Performance of a High-Throughput Screen for Anti-Mycobacterial Compounds”. In: *PloS One* 9.5 (2014), e96348. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0096348. pmid: 24788852.

[20] S. H. Choudhri et al. “Pharmacokinetics of Antimycobacterial Drugs in Patients with Tuberculosis, AIDS, and Diarrhea”. In: *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 25.1 (1997), pp. 104–111. ISSN: 1058-4838. pmid: 9243044.

[21] Corey L. Compton et al. “Antibacterial Activity of and Resistance to Small Molecule Inhibitors of the ClpP Peptidase”. In: *ACS chemical biology* 8.12 (20, 2013), pp. 2669–2677. ISSN: 1554-8937. DOI: 10.1021/cb400577b. pmid: 24047344.

[22] Véronique Dartois. “The Path of Anti-Tuberculosis Drugs: From Blood to Lesions to Mycobacterial Cells”. In: *Nature Reviews. Microbiology* 12.3 (2014), pp. 159–167. ISSN: 1740-1534. DOI: 10.1038/nrmicro3200. pmid: 24487820.

[23] Michael A DeJesus and Thomas R Ioerger. “A Hidden Markov Model for Identifying Essential and Growth-Defect Regions in Bacterial Genomes from Transposon Insertion Sequencing Data”. In: *BMC Bioinformatics* 14.1 (2013). ISSN: 1471-2105. DOI: 10.1186/1471-2105-14-303. URL: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-303 (visited on 02/03/2019).

[24] Michael A. DeJesus et al. “Comprehensive Essentiality Analysis of the Mycobacterium Tuberculosis Genome via Saturating Transposon Mutagenesis”. In: *mBio* 8.1 (17, 2017). ISSN: 2150-7511. DOI: 10.1128/mBio.02133-16. pmid: 28096490.

[25] Michael A. DeJesus et al. “TRANSIT–A Software Tool for Himar1 TnSeq Analysis”. In: *PLoS computational biology* 11.10 (2015), e1004401. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004401. pmid: 26447887.

[26] Peter R. Donald and Andreas H. Diacon. "Para-Aminosalicylic Acid: The Return of an Old Friend". In: *The Lancet. Infectious Diseases* 15.9 (2015), pp. 1091–1099. ISSN: 1474-4457. DOI: 10.1016/S1473-3099(15)00263-7. pmid: 26277036.

[27] Kelly E. Dooley and Richard E. Chaisson. "Tuberculosis and Diabetes Mellitus: Convergence of Two Epidemics". In: *The Lancet. Infectious Diseases* 9.12 (2009), pp. 737–746. ISSN: 1474-4457. DOI: 10.1016/S1473-3099(09)70282-8. pmid: 19926034.

[28] R. J. Dubos and B. D. Davis. "Factors Affecting the Growth of Tubercle Bacilli in Liquid Media". In: *The Journal of Experimental Medicine* 83 (1946), pp. 409–423. ISSN: 0022-1007. pmid: 21025659.

[29] Joseph O. Falkinham. "Nontuberculous Mycobacteria from Household Plumbing of Patients with Nontuberculous Mycobacteria Disease". In: *Emerging Infectious Diseases* 17.3 (2011), pp. 419–424. ISSN: 1080-6059. DOI: 10.3201/eid1703.101510. pmid: 21392432.

[30] M. J. Ferrándiz et al. "Reactive Oxygen Species Contribute to the Bactericidal Effects of the Fluoroquinolone Moxifloxacin in Streptococcus Pneumoniae". In: *Antimicrobial Agents and Chemotherapy* 60.1 (2016), pp. 409–417. ISSN: 1098-6596. DOI: 10.1128/AAC.02299-15. pmid: 26525786.

[31] D. N. Fish et al. "Penetration of Clarithromycin into Lung Tissues from Patients Undergoing Lung Resection". In: *Antimicrobial Agents and Chemotherapy* 38.4 (1994), pp. 876–878. ISSN: 0066-4804. pmid: 8031063.

[32] R. Andres Floto et al. "US Cystic Fibrosis Foundation and European Cystic Fibrosis Society Consensus Recommendations for the Management of Non-Tuberculous Mycobacteria in Individuals with Cystic Fibrosis". In: *Thorax* 71 Suppl 1 (2016), pp. i1–22. ISSN: 1468-3296. DOI: 10.1136/thoraxjnl-2015-207360. pmid: 26666259.

[33] E. M. Foley-Thomas et al. "Phage Infection, Transfection and Transformation of Mycobacterium Avium Complex and Mycobacterium Paratuberculosis". In: *Microbiology (Reading, England)* 141 (Pt 5) (1995), pp. 1173–1181. ISSN: 1350-0872. DOI: 10.1099/13500872-141-5-1173. pmid: 7773411.

[34] The Bill and Melinda Gates Foundation. *Tuberculosis Strategy Overview*. 2012. URL: https://docs.gatesfoundation.org/documents/tuberculosis-strategy.pdf.

[35] W. Fox, G. A. Ellard, and D. A. Mitchison. "Studies on the Treatment of Tuberculosis Undertaken by the British Medical Research Council Tuberculosis Units, 1946-1986, with Relevant Subsequent Publications". In: *The International Journal of Tuberculosis and Lung Disease: The Official Journal of the International Union Against Tuberculosis and Lung Disease* 3 (10 Suppl 2 1999), S231–279. ISSN: 1027-3719. pmid: 10529902.

[36] Martin Gengenbacher et al. "Nutrient-Starved, Non-Replicating Mycobacterium Tuberculosis Requires Respiration, ATP Synthase and Isocitrate Lyase for Maintenance of ATP Homeostasis and Viability". In: *Microbiology (Reading, England)* 156 (Pt 1 2010), pp. 81–87. ISSN: 1465-2080. DOI: 10.1099/mic.0.033084-0. pmid: 19797356.

[37] Stephen H. Gillespie et al. "Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis". In: *New England Journal of Medicine* 371.17 (23, 2014), pp. 1577–1587. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa1407426. URL: http://www.nejm.org/doi/10.1056/NEJMoa1407426 (visited on 04/18/2019).

[38] Chunling Gong et al. "Biochemical and Genetic Analysis of the Four DNA Ligases of Mycobacteria". In: *The Journal of Biological Chemistry* 279.20 (14, 2004), pp. 20594–20606. ISSN: 0021-9258. DOI: 10.1074/jbc.M401841200. pmid: 14985346.

[39] Andrew L. Goodman et al. "Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat". In: *Cell Host & Microbe* 6.3 (17, 2009), pp. 279–289. ISSN: 1934-6069. DOI: 10.1016/j.chom.2009.08.003. pmid: 19748469.

[40] Sarah Schmidt Grant et al. “Identification of Novel Inhibitors of Nonreplicating Mycobacterium Tuberculosis Using a Carbon Starvation Model”. In: *ACS chemical biology* 8.10 (18, 2013), pp. 2224–2234. ISSN: 1554-8937. DOI: 10.1021/cb4004817. pmid: 23898841.

[41] David E. Griffith et al. “An Official ATS/IDSA Statement: Diagnosis, Treatment, and Prevention of Nontuberculous Mycobacterial Diseases”. In: *American Journal of Respiratory and Critical Care Medicine* 175.4 (15, 2007), pp. 367–416. ISSN: 1073-449X. DOI: 10.1164/rccm.200604-571ST. pmid: 17277290.

[42] David E. Griffith et al. “Semiquantitative Culture Analysis during Therapy for Mycobacterium Avium Complex Lung Disease”. In: *American Journal of Respiratory and Critical Care Medicine* 192.6 (15, 2015), pp. 754–760. ISSN: 1535-4970. DOI: 10.1164/rccm.201503-04440C. pmid: 26068042.

[43] Shashank Gupta et al. “Acceleration of Tuberculosis Treatment by Adjunctive Therapy with Verapamil as an Efflux Inhibitor”. In: *American Journal of Respiratory and Critical Care Medicine* 188.5 (1, 2013), pp. 600–607. ISSN: 1535-4970. DOI: 10.1164/rccm.201304-06500C. pmid: 23805786.

[44] Samuel Halstrom, Patricia Price, and Rachel Thomson. “Review: Environmental Mycobacteria as a Cause of Human Infection”. In: *International Journal of Mycobacteriology* 4.2 (2015), pp. 81–91. ISSN: 22125531. DOI: 10.1016/j.ijmyco.2015.03.002. URL: http://www.ijmyco.org/article.asp?issn=2212-5531;year=2015;volume=4;issue=2;spage=81;epage=91;aulast=Halstrom (visited on 03/28/2019).

[45] Victoria Hansen et al. “Infectious Disease Mortality Trends in the United States, 1980-2014”. In: *JAMA* 316.20 (22, 2016), pp. 2149–2151. ISSN: 1538-3598. DOI: 10.1001/jama.2016.12423. pmid: 27893119.

[46] Melanie J. Harriff et al. “Mycobacterium Avium Genes MAV_5138 and MAV_3679 Are Transcriptional Regulators That Play a Role in Invasion of Epithelial Cells, in Part by Their Regulation of CipA, a Putative Surface Protein Interacting with Host Cell Signaling Pathways”. In: *Journal of Bacteriology* 191.4 (2009), pp. 1132–1142. ISSN: 1098-5530. DOI: 10.1128/JB.01359-07. pmid: 19060135.

[47] Wouter Hoefsloot et al. “The Geographic Diversity of Nontuberculous Mycobacteria Isolated from Pulmonary Samples: An NTM-NET Collaborative Study”. In: *The European Respiratory Journal* 42.6 (2013), pp. 1604–1613. ISSN: 1399-3003. DOI: 10.1183/09031936.00149212. pmid: 23598956.

[48] Rein M. G. J. Houben and Peter J. Dodd. “The Global Burden of Latent Tuberculosis Infection: A Re-Estimation Using Mathematical Modelling”. In: *PLoS medicine* 13.10 (2016), e1002152. ISSN: 1549-1676. DOI: 10.1371/journal.pmed.1002152. pmid: 27780211.

[49] United States White House. *National Action Plan for Combating Multidrug-Resistant Tuberculosis*. 2015.

[50] L. Hudek et al. “Role of Phosphate Transport System Component PstB1 in Phosphate Internalization by Nostoc Punctiforme”. In: *Applied and Environmental Microbiology* 82.21 (1, 2016). Ed. by S.-J. Liu, pp. 6344–6356. ISSN: 0099-2240, 1098-5336. DOI: 10.1128/AEM.01336-16. URL: http://aem.asm.org/lookup/doi/10.1128/AEM.01336-16 (visited on 04/17/2019).

[51] Michael D. Iseman. *A Clinician’s Guide to Tuberculosis*. Philadelphia: Lippincott Williams & Wilkins, 2000. 460 pp. ISBN: 978-0-7817-1749-6.

[52] Zeyaul Islam et al. “Structural Basis for Competitive Inhibition of 3,4-Dihydroxy-2-Butanone-4-Phosphate Synthase from Vibrio Cholerae”. In: *The Journal of Biological Chemistry* 290.18 (1, 2015), pp. 11293–11308. ISSN: 1083-351X. DOI: 10.1074/jbc.M114.611830. pmid: 25792735.

[53] J.-P. Janssens and H. L. Rieder. "An Ecological Analysis of Incidence of Tuberculosis and per Capita Gross Domestic Product". In: *The European Respiratory Journal* 32.5 (2008), pp. 1415–1416. ISSN: 1399-3003. DOI: 10.1183/09031936.00078708. pmid: 18978146.

[54] Jinshan Jin et al. "SecA Inhibitors as Potential Antimicrobial Agents: Differential Actions on SecA-Only and SecA-SecYEG Protein-Conducting Channels". In: *FEMS microbiology letters* 365.15 (1, 2018). ISSN: 1574-6968. DOI: 10.1093/femsle/fny145. pmid: 30007321.

[55] A. R. Jonckheere. "A Distribution-Free k-Sample Test Against Ordered Alternatives". In: *Biometrika* 41.1/2 (1954), p. 133. ISSN: 00063444. DOI: 10.2307/2333011. URL: https://www.jstor.org/stable/2333011?origin=crossref (visited on 04/08/2019).

[56] Brian A. Kendall and Kevin L. Winthrop. "Update on the Epidemiology of Pulmonary Nontuberculous Mycobacterial Infections". In: *Seminars in Respiratory and Critical Care Medicine* 34.1 (2013), pp. 87–94. ISSN: 1098-9048. DOI: 10.1055/s-0033-1333567. pmid: 23460008.

[57] Joses Muthuri Kirigia and Rosenabi Deborah Karimi Muthuri. "Productivity Losses Associated with Tuberculosis Deaths in the World Health Organization African Region". In: *Infectious Diseases of Poverty* 5.1 (1, 2016), p. 43. ISSN: 2049-9957. DOI: 10.1186/s40249-016-0138-5. pmid: 27245156.

[58] Won-Jung Koh et al. "Treatment of Refractory Mycobacterium Avium Complex Lung Disease with a Moxifloxacin-Containing Regimen". In: *Antimicrobial Agents and Chemotherapy* 57.5 (2013), pp. 2281–2285. ISSN: 1098-6596. DOI: 10.1128/AAC.02281-12. pmid: 23478956.

[59] Sergey Koren et al. "Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation". In: *Genome Research* 27.5 (2017), pp. 722–736. ISSN: 1549-5469. DOI: 10.1101/gr.215087.116. pmid: 28298431.

[60] Ekaterina Kulchavenya. "Extrapulmonary Tuberculosis: Are Statistical Reports Accurate?" In: *Therapeutic Advances in Infectious Disease* 2.2 (2014), pp. 61–70. ISSN: 2049-9361. DOI: 10.1177/2049936114528173. pmid: 25165556.

[61] Trent Kunkle et al. "Hydroxybiphenylamide GroEL/ES Inhibitors Are Potent Antibacterials against Planktonic and Biofilm Forms of Staphylococcus Aureus". In: *Journal of Medicinal Chemistry* (15, 2018). ISSN: 1520-4804. DOI: 10.1021/acs.jmedchem.8b01293. pmid: 30392371.

[62] Ben Langmead and Steven L. Salzberg. "Fast Gapped-Read Alignment with Bowtie 2". In: *Nature Methods* 9.4 (4, 2012), pp. 357–359. ISSN: 1548-7105. DOI: 10.1038/nmeth.1923. pmid: 22388286.

[63] Gemma C. Langridge et al. "Simultaneous Assay of Every Salmonella Typhi Gene Using One Million Transposon Mutants". In: *Genome Research* 19.12 (2009), pp. 2308–2316. ISSN: 1549-5469. DOI: 10.1101/gr.097097.109. pmid: 19826075.

[64] Heng Li et al. "The Sequence Alignment/Map Format and SAMtools". In: *Bioinformatics (Oxford, England)* 25.16 (15, 2009), pp. 2078–2079. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btp352. pmid: 19505943.

[65] Yongjun Li et al. "A Mycobacterium Avium PPE Gene Is Associated with the Ability of the Bacterium to Grow in Macrophages and Virulence in Mice". In: *Cellular Microbiology* 7.4 (2005), pp. 539–548. ISSN: 1462-5814. DOI: 10.1111/j.1462-5822.2004.00484.x. pmid: 15760454.

[66] R. H. Liss, R. J. Letourneau, and J. P. Schepis. "Distribution of Ethambutol in Primate Tissues and Cells". In: *The American Review of Respiratory Disease* 123.5 (1981), pp. 529–532. ISSN: 0003-0805. DOI: 10.1164/arrd.1981.123.5.529. pmid: 6786149.

[67] J. K. Lodge, K. Weston-Hafer, and D. E. Berg. "Transposon Tn5 Target Specificity: Preference for Insertion at G/C Pairs". In: *Genetics* 120.3 (1988), pp. 645–650. ISSN: 0016-6731. pmid: 2852135.

[68] R. O. Loebel, E. Shorr, and H. B. Richardson. "The Influence of Adverse Conditions upon the Respiratory Metabolism and Growth of Human Tubercle Bacilli". In: *Journal of Bacteriology* 26.2 (1933), pp. 167–200. ISSN: 0021-9193. pmid: 16559650.

[69] Jarukit E. Long et al. "Identifying Essential Genes in Mycobacterium Tuberculosis by Global Phenotypic Profiling". In: *Methods in Molecular Biology (Clifton, N.J.)* 1279 (2015), pp. 79–95. ISSN: 1940-6029. DOI: 10.1007/978-1-4939-2398-4_6. pmid: 25636614.

[70] Shichun Lun and William R. Bishai. "Characterization of a Novel Cell Wall-Anchored Protein with Carboxylesterase Activity Required for Virulence in Mycobacterium Tuberculosis". In: *The Journal of Biological Chemistry* 282.25 (22, 2007), pp. 18348–18356. ISSN: 0021-9258. DOI: 10.1074/jbc.M700035200. pmid: 17428787.

[71] Gaurav Majumdar et al. "Genome-Wide Transposon Mutagenesis in Mycobacterium Tuberculosis and Mycobacterium Smegmatis". In: *Methods in Molecular Biology (Clifton, N.J.)* 1498 (2017), pp. 321–335. ISSN: 1940-6029. DOI: 10.1007/978-1-4939-6472-7_21. pmid: 27709585.

[72] Lionel A. Mandell et al. "Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults". In: *Clinical Infectious Diseases* 44 (Supplement_2 1, 2007), S27–S72. ISSN: 1537-6591, 1058-4838. DOI: 10.1086/511159. URL: http://academic.oup.com/cid/article/44/Supplement_2/S27/372079/Infectious-Diseases-Society-of-AmericaAmerican (visited on 03/28/2019).

[73] Guillaume Marçais et al. "MUMmer4: A Fast and Versatile Genome Alignment System". In: *PLoS computational biology* 14.1 (2018), e1005944. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005944. pmid: 29373581.

[74] William M. Matern, Joel S. Bader, and Petros C. Karakousis. "Genome Analysis of Mycobacterium Avium Subspecies Hominissuis Strain 109". In: *Scientific Data* 5 (4, 2018), p. 180277. ISSN: 2052-4463. DOI: 10.1038/sdata.2018.277. pmid: 30512015.

[75] M. McNabe et al. "Mycobacterium Avium Ssp. Hominissuis Biofilm Is Composed of Distinct Phenotypes and Influenced by the Presence of Antimicrobials". In: *Clinical Microbiology and Infection* 17.5 (2011), pp. 697–703. ISSN: 1198743X. DOI: 10.1111/j.1469-0691.2010.03307.x. URL: https://linkinghub.elsevier.com/retrieve/pii/S1198743X15602610 (visited on 03/28/2019).

[76] Kenan C. Murphy et al. "ORBIT: A New Paradigm for Genetic Engineering of Mycobacterial Chromosomes". In: *mBio* 9.6 (12, 2018). ISSN: 2150-7511. DOI: 10.1128/mBio.01467-18. pmid: 30538179.

[77] John F. Murray, Dean E. Schraufnagel, and Philip C. Hopewell. "Treatment of Tuberculosis. A Historical Perspective". In: *Annals of the American Thoracic Society* 12.12 (2015), pp. 1749–1759. ISSN: 2325-6621. DOI: 10.1513/AnnalsATS.201509-632PS. pmid: 26653188.

[78] Jacqueline L. Naffin-Olivos et al. "Mycobacterium Tuberculosis Hip1 Modulates Macrophage Responses through Proteolysis of GroEL2". In: *PLoS pathogens* 10.5 (2014), e1004132. ISSN: 1553-7374. DOI: 10.1371/journal.ppat.1004132. pmid: 24830429.

[79] G. L. Newton, N. Buchmeier, and R. C. Fahey. "Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of Actinobacteria". In: *Microbiology and Molecular Biology Reviews* 72.3 (1, 2008), pp. 471–494. ISSN: 1092-2172. DOI: 10.1128/MMBR.00008-08. URL: http://mmbr.asm.org/cgi/doi/10.1128/MMBR.00008-08 (visited on 04/17/2019).

[80] J. O'Grady et al. "Tuberculosis in Prisons: Anatomy of Global Neglect". In: *The European Respiratory Journal* 38.4 (2011), pp. 752–754. ISSN: 1399-3003. DOI: 10.1183/09031936.00041211. pmid: 21965498.

[81] Anil K. Ojha, William R. Jacobs, and Graham F. Hatfull. "Genetic Dissection of Mycobacterial Biofilms". In: *Methods in Molecular Biology (Clifton, N.J.)* 1285 (2015), pp. 215–226. ISSN: 1940-6029. DOI: 10.1007/978-1-4939-2450-9_12. pmid: 25779318.

[82] D. Rebecca Prevots and Theodore K. Marras. "Epidemiology of Human Pulmonary Infection with Nontuberculous Mycobacteria: A Review". In: *Clinics in Chest Medicine* 36.1 (2015), pp. 13–34. ISSN: 1557-8216. DOI: 10.1016/j.ccm.2014.10.002. pmid: 25676516.

[83] Brendan Prideaux et al. "The Association between Sterilizing Activity and Drug Distribution into Tuberculosis Lesions". In: *Nature Medicine* 21.10 (2015), pp. 1223–1227. ISSN: 1546-170X. DOI: 10.1038/nm.3937. pmid: 26343800.

[84] R. P. Prioli, A. Tanna, and I. N. Brown. "Rapid Methods for Counting Mycobacteria–Comparison of Methods for Extraction of Mycobacterial Adenosine Triphosphate (ATP) Determined by Firefly Luciferase Assay". In: *Tubercle* 66.2 (1985), pp. 99–108. ISSN: 0041-3879. pmid: 3895682.

[85] Jeremy M. Rock et al. "Programmable Transcriptional Repression in Mycobacteria Using an Orthogonal CRISPR Interference Platform". In: *Nature Microbiology* 2 (6, 2017), p. 16274. ISSN: 2058-5276. DOI: 10.1038/nmicrobiol.2016.274. pmid: 28165460.

[86] E. J. Rubin et al. "In Vivo Transposition of Mariner-Based Elements in Enteric Bacteria and Mycobacteria". In: *Proceedings of the National Academy of Sciences of the United States of America* 96.4 (16, 1999), pp. 1645–1650. ISSN: 0027-8424. pmid: 9990078.

[87] C. M. Sassetti, D. H. Boyd, and E. J. Rubin. "Comprehensive Identification of Conditionally Essential Genes in Mycobacteria". In: *Proceedings of the National Academy of Sciences of the United States of America* 98.22 (23, 2001), pp. 12712–12717. ISSN: 0027-8424. DOI: 10.1073/pnas.231275498. pmid: 11606763.

[88] Kazunori Sawada et al. "Effect of Pyruvate Kinase Gene Deletion on the Physiology of Corynebacterium Glutamicum ATCC13032 under Biotin-Sufficient Non-Glutamate-Producing Conditions: Enhanced Biomass Production". In: *Metabolic Engineering Communications* 2 (2015), pp. 67–75. ISSN: 22140301. DOI: 10.1016/j.meteno.2015.07.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S2214030115300067 (visited on 04/18/2019).

[89] Sung Jae Shin et al. "Efficient Differentiation of Mycobacterium Avium Complex Species and Subspecies by Use of Five-Target Multiplex PCR". In: *Journal of Clinical Microbiology* 48.11 (2010), pp. 4057–4062. ISSN: 1098-660X. DOI: 10.1128/JCM.00904-10. pmid: 20810779.

[90] Yun Su Sim et al. "Standardized Combination Antibiotic Treatment of Mycobacterium Avium Complex Lung Disease". In: *Yonsei Medical Journal* 51.6 (2010), pp. 888–894. ISSN: 1976-2437. DOI: 10.3349/ymj.2010.51.6.888. pmid: 20879056.

[91] Sara E. Strollo et al. "The Burden of Pulmonary Nontuberculous Mycobacterial Disease in the United States". In: *Annals of the American Thoracic Society* 12.10 (2015), pp. 1458–1464. ISSN: 2325-6621. DOI: 10.1513/AnnalsATS.201503-173OC. pmid: 26214350.

[92] David Stuckler et al. "Mining and Risk of Tuberculosis in Sub-Saharan Africa". In: *American Journal of Public Health* 101.3 (2011), pp. 524–530. ISSN: 1541-0048. DOI: 10.2105/AJPH.2009.175646. pmid: 20516372.

[93] Susan Swindells et al. "One Month of Rifapentine plus Isoniazid to Prevent HIV-Related Tuberculosis". In: *The New England Journal of Medicine* 380.11 (14, 2019), pp. 1001–1011. ISSN: 1533-4406. DOI: 10.1056/NEJMoa1806808. pmid: 30865794.

[94] Amish Talwar et al. "Tuberculosis — United States, 2018". In: *MMWR. Morbidity and Mortality Weekly Report* 68.11 (22, 2019), pp. 257–262. ISSN: 0149-2195, 1545-861X. DOI: 10.15585/mmwr.mm6811a2. URL: http://www.cdc.gov/mmwr/volumes/68/wr/mm6811a2.htm?s_cid=mm6811a2_w (visited on 04/03/2019).

[95] Mai Ping Tan et al. “Nitrate Respiration Protects Hypoxic Mycobacterium Tuberculosis against Acid- and Reactive Nitrogen Species Stresses”. In: *PloS One* 5.10 (26, 2010), e13356. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0013356. pmid: 21048946.

[96] Tatiana Tatusova et al. “NCBI Prokaryotic Genome Annotation Pipeline”. In: *Nucleic Acids Research* 44.14 (19, 2016), pp. 6614–6624. ISSN: 1362-4962. DOI: 10.1093/nar/gkw569. pmid: 27342282.

[97] T.J. Terpstra. “The Asymptotic Normality and Consistency of Kendall’s Test against Trend, When Ties Are Present in One Ranking.” In: *Indigationes Mathematicae* 14.3 (1952), pp. 327–333.

[98] Petri Törönen, Alan Medlar, and Liisa Holm. “PANNZER2: A Rapid Functional Annotation Web Server”. In: *Nucleic Acids Research* 46 (8, 2018), W84–W88. ISSN: 1362-4962. DOI: 10.1093/nar/gky350. pmid: 29741643.

[99] Kei-Ichi Uchiya et al. “Comparative Genome Analyses of Mycobacterium Avium Reveal Genomic Features of Its Subspecies and Strains That Cause Progression of Pulmonary Disease”. In: *Scientific Reports* 7 (3, 2017), p. 39750. ISSN: 2045-2322. DOI: 10.1038/srep39750. pmid: 28045086.

[100] J. van Ingen et al. “Clinical Relevance of Non-Tuberculous Mycobacteria Isolated in the Nijmegen-Arnhem Region, The Netherlands”. In: *Thorax* 64.6 (2009), pp. 502–506. ISSN: 1468-3296. DOI: 10.1136/thx.2008.110957. pmid: 19213773.

[101] Jakko van Ingen et al. “The Pharmacokinetics and Pharmacodynamics of Pulmonary Mycobacterium Avium Complex Disease Treatment”. In: *American Journal of Respiratory and Critical Care Medicine* 186.6 (15, 2012), pp. 559–565. ISSN: 1535-4970. DOI: 10.1164/rccm.201204-0682OC. pmid: 22744719.

[102] Heleen Van Acker and Tom Coenye. “The Role of Reactive Oxygen Species in Antibiotic-Mediated Killing of Bacteria”. In: *Trends in Microbiology* 25.6 (2017), pp. 456–466. ISSN: 1878-4380. DOI: 10.1016/j.tim.2016.12.008. pmid: 28089288.

[103] Robert Vaser et al. “Fast and Accurate de Novo Genome Assembly from Long Uncorrected Reads”. In: *Genome Research* 27.5 (2017), pp. 737–746. ISSN: 1549-5469. DOI: 10.1101/gr.214270.116. pmid: 28100585.

[104] Laura E. Via et al. “Tuberculous Granulomas Are Hypoxic in Guinea Pigs, Rabbits, and Nonhuman Primates”. In: *Infection and Immunity* 76.6 (2008), pp. 2333–2340. ISSN: 1098-5522. DOI: 10.1128/IAI.01515-07. pmid: 18347040.

[105] Catherine Vilchèze et al. “Enhanced Respiration Prevents Drug Tolerance and Drug Resistance in Mycobacterium Tuberculosis”. In: *Proceedings of the National Academy of Sciences of the United States of America* 114.17 (25, 2017), pp. 4495–4500. ISSN: 1091-6490. DOI: 10.1073/pnas.1704376114. pmid: 28396391.

[106] Yuichi Wakamoto et al. “Dynamic Persistence of Antibiotic-Stressed Mycobacteria”. In: *Science (New York, N.Y.)* 339.6115 (4, 2013), pp. 91–95. ISSN: 1095-9203. DOI: 10.1126/science.1229858. pmid: 23288538.

[107] Bruce J. Walker et al. “Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement”. In: *PloS One* 9.11 (2014), e112963. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0112963. pmid: 25409509.

[108] WHO. *Priorities for Tuberculosis Research*. 2013. URL: https://www.who.int/tdr/publications/tuberculosis_research/en/.

[109] Ryan R. Wick et al. “Bandage: Interactive Visualization of de Novo Genome Assemblies”. In: *Bioinformatics (Oxford, England)* 31.20 (15, 2015), pp. 3350–3352. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btv383. pmid: 26099265.

[110] Ryan R. Wick et al. “Unicycler: Resolving Bacterial Genome Assemblies from Short and Long Sequencing Reads”. In: *PLoS computational biology* 13.6 (2017), e1005595. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005595. pmid: 28594827.

[111] World Health Organization. *Global Tuberculosis Report 2018*. 2018. ISBN: 978-92-4-156564-6.

[112] Zhifang Xie, Noman Siddiqi, and Eric J. Rubin. “Differential Antibiotic Susceptibilities of Starved Mycobacterium Tuberculosis Isolates”. In: *Antimicrobial Agents and Chemotherapy* 49.11 (2005), pp. 4778–4780. ISSN: 0066-4804. DOI: 10.1128/AAC.49.11.4778-4780.2005. pmid: 16251329.

[113] Michelle Yee et al. “Draft Genome Sequence of Mycobacterium Avium 11”. In: *Genome Announcements* 5.32 (10, 2017). ISSN: 2169-8287. DOI: 10.1128/genomeA.00766-17. pmid: 28798178.

[114] D. V. Zaykin et al. “Truncated Product Method for Combining P-Values”. In: *Genetic Epidemiology* 22.2 (2002), pp. 170–185. ISSN: 0741-0395. DOI: 10.1002/gepi.0042. pmid: 11788962.

[115] Ying Zhang. “Persisters, Persistent Infections and the Yin-Yang Model”. In: *Emerging Microbes & Infections* 3.1 (2014), e3. ISSN: 2222-1751. DOI: 10.1038/emi.2014.3. pmid: 26038493.

WILLIAM MATERN

Email: wmatern1@jhmi.edu | Tel: 603-382-6653 | Baltimore, MD

EDUCATION

Johns Hopkins University School of Medicine, Baltimore, MD
Doctor of Philosophy, Biomedical Engineering
Aug. 2012 - Present
Graduation: May 2019 (Available: August 2019)

University of New Hampshire, Durham, NH
Bachelor of Science, Mechanical Engineering, *summa cum laude* (3.75/4.00)
Minor in Applied Mathematics
Sept. 2009 - May 2012

EXPERIENCE

PhD Student — Dr. Joel Bader (Advisor), Dr. Petros Karakousis (Advisor), Baltimore, MD — Jan. 2013 - Present

- Thesis title: "A systems biology approach to identify adjunctive drug targets in two major mycobacterial pathogens".
- Developed a combined experimental/computational approach to identify enzymes that might be targeted to achieve cure faster for *Mycobacterium tuberculosis* and *Mycobacterium avium*.
- Optimized custom preps for next generation sequencing (no kit available)
- Identified the essential genes of *Mycobacterium avium*. A manuscript is in preparation.
- Combined non-parametric statistics, multiple hypothesis correction, and maximum likelihood estimation into a processing pipeline to identify mutants with defects from sequence data.
- Validated mycobacterial mutants as hypersusceptible to currently used drugs. A manuscript is planned for reporting the results from this work.

Teaching Assistant — Dr. Joel Bader, Baltimore, MD — Sept. - Dec. 2014

- Systems Bioengineering III (Dept. of Biomedical Engineering).
- Prepared a review section once a week, made handouts, wrote homework solutions, administered and graded tests.

Student Researcher — Bob Robey, Los Alamos, NM — Summer 2012

- Los Alamos National Lab in the Computational Physics Division (XCP)
- Compared methods for modelling discontinuities in the solution of partial differential equations with a focus on finite element methods.

Undergraduate Research — Dr. Greg Chini, Durham, NH — Fall 2010 - Spring 2012

- Computational models of fluid-solid interactions in mammalian lungs leading to a publication in the UNH undergraduate research journal: `http://scholars.unh.edu/inquiry_2012/`.

COMPUTER AND ANALYSIS SKILLS

- Analysis of Illumina (RNA-seq, Tn-seq, DNA-seq, etc.) and PacBio data.
- Languages: Proficient with Python, MATLAB; Experienced with bash, GNU Make, Latex, C, R
- Operating Systems: Linux (Arch, Ubuntu), Windows
- Packages: NumPy, SciPy, Matplotlib, Pandas, DESeq2, Bowtie2, BWA, Mummer, Spades.
- Application of parametric and non-parametric statistical methods in a research setting to sequence data.
- Machine learning coursework (graphical models (HMMs), SVMs, stochastic optimization, basic NNs, etc.)
- Flux balance analysis
- Exposure to parallelization and high-performance computing.
- Implementation of numerial PDE solvers (e.g. Finite Difference and Finite Elements)
- Github (scientific projects): `https://github.com/joelbader/fast_allvsall_scripts/tree/v1.1`, `https://github.com/joelbader/regulon-enrichment`

LABORATORY SKILLS

- Protocol optimization of library preps for next generation sequencing (Illumina).
- High-throughput genetic screens via transposon mutagenesis (Tn-seq)
- Proficient in the safe handling of BSL-3 organisms (*Mycobacterium tuberculosis*)
- Set up, repair, and use of an anaerobic chamber
- Molecular biology: primer design, RT-qPCR, gel electrophoresis, Southern blots, cloning, bacteriophage
- DNA and RNA extraction from bacterial cells.

PEER-REVIEWED PUBLICATIONS

- **Matern, W. M.**, Bader, J. S., & Karakousis, P. C. (2018). *Genome analysis of Mycobacterium avium subspecies hominissuis strain 109.* Scientific Data. Vol. 5. `http://doi.org/10.1038/sdata.2018.277`
- **Matern, W. M.**, Rifat, D., Bader, J. S., & Karakousis, P. C. (2018). *Gene Enrichment Analysis Reveals Major Regulators of Mycobacterium tuberculosis Gene Expression in Two Models of Antibiotic Tolerance.* Frontiers in Microbiology Vol. 9, p610. `http://doi.org/10.3389/fmicb.2018.00610`
- Pienaar, E., **Matern, W. M.**, Linderman, J. J., Bader, J. S., & Kirschner, D. E. (2016). *Multiscale model of Mycobacterium tuberculosis infection maps metabolite and gene perturbations to granuloma sterilization predictions.* Infection and Immunity, 84(5), 1650–1669. `http://doi.org/10.1128/IAI.01438-15`

POSTER PRESENTATIONS

- Matern, W.M., Ioerger, T.R., Bader, J.S., Karakousis, P.C. *Finding Antibiotic Persistence Genes in Mycobacteria with Transposon Sequencing.* University of Maryland. 6/4/18. Baltimore, MD.
- Matern, W.M., Ioerger, T.R., Bader, J.S., Karakousis, P.C. *Finding Antibiotic Persistence Genes in Mycobacteria with Transposon Sequencing.* Keystone Symposium Conference - Tuberculosis: Translating Scientific Findings for Clinical and Public Health Impact. 4/11/18. Whistler, British Columbia.
- Matern, W.M., Rifat, D., Ioerger, T.R., Karakousis, P.C., Bader, J.S. *Identifying major regulators of antibiotic tolerance in Mycobacterium tuberculosis and avium.* Johns Hopkins University Department of Medicine Retreat. 3/16/18. Baltimore, MD.

ORAL PRESENTATIONS

- Matern, W.M., Bader, J.S., Karakousis, P.C. *Identifying Adjunctive Mycobacterial Targets for Treatment Shortening.* 15 minute talk. Johns Hopkins University Center for Tuberculosis Research Annual Scientific Meeting. 6/12/18. Baltimore, MD.
- Matern, W.M., Rifat, D., Bader, J.S., Karakousis, P.C. *Mtb dormancy and resuscitation: A systems biology-based approach.* 15 minute talk. Johns Hopkins University Center for Tuberculosis Research Annual Scientific Meeting. 6/10/15. Baltimore, MD.

- Matern, W.M., Pienaar, E., Bader, J.S., Kirschner, D.E. *A multi-scale model of bacterial metabolism and macrophage dynamics to better understand host-pathogen interactions during TB latency.* 15 minute summary of results from research in progress. Tb Systems Biology Investigator's Meeting. National Heart, Lung and Blood Institute. 9/12/14. Bethesda, MD.
- Matern, W.M., Bader, J.S. *Multiscale Models of Mycobacterium Tuberculosis.* 15 minute presentation of ongoing research. Johns Hopkins Institute of Computational Medicine Retreat. 5/19/14. Baltimore, MD.

RESEARCH PROJECTS FUNDED

- Identifying the mechanism of tolerance to rifampin mediated by caeA/Hip1/Rv2224c in Mycobacterium tuberculosis.
  Sponsor: Potts Memorial Foundation
  Period: Jan. 2019 - Dec. 2019.
- Identifying molecular targets for preventing multidrug tolerance in Mycobacterium avium infection.
  Sponsor: Sherrilyn and Ken Fisher Center for Environmental Infectious Diseases
  Period: Jan. 2016 - Jan. 2017.
- An investigation of mathematical modeling techniques as applied to thin films in an alveolar network.
  Sponsor: University of New Hampshire Hamel Center - Summer Undergraduate Research Fellowship
  Period: Summer 2011.

HONORS AND AWARDS

- The Arthur M. Dannenberg Jr. Award for Tuberculosis Research (2018)
- NSF Fellowship to attend the 13th Annual International Summer School on Biocomplexity and Biodesign at Bogazici University (2014)
- University of New Hampshire - Tau Beta Pi Member
- University of New Hampshire - Presidential Scholarship (2009 - 2012)
- University of New Hampshire - Kenneth J. Higson Scholarship (2011)
- University of New Hampshire - University Honors Program

STUDENTS MENTORED

- Robert Jenquin (Volunteer) - June 2018 - Pres.
- Leah Hoover (Undergrad) - Feb. 2017 - May 2018
- Grace Ren (Undergrad) - Mar. 2017 - Sept. 2018
- Sameer Thakker (Undergrad) - Summer 2016
- Christina Blonski (PostBac) - Summer 2016
- Jennifer Mendez (High School) - Summer 2015 - Summer 2016

COMMUNITY

- Thread Inc. (previous name: Incentive Mentoring Program) Mentor + Leader - I provide advice and support for a young adult growing up in Baltimore city (Summer 2013 - Pres.)
- University of New Hampshire - Tutor in Mechanical Engineering
- University of New Hampshire - Tutor for Tau Beta Pi