The Effects of Functional Variability of *M. smegmatis* on Rifampicin Tolerance

A Thesis Submitted By

Kirill Richardson

In Partial Fulfillment of the Requirements

for the Degree of Master of Science

in Biomedical Engineering

TUFTS UNIVERSITY

August 2015

Advisor: Bree Aldridge, PhD.

Abstract

Tuberculosis remains the world's second deadliest infectious disease after HIV/AIDS, and poses a major threat to public health in many parts of the world. It has previously been demonstrated that mycobacterial colonies exhibit distinct heterogeneous phenotypes, and that this functional variation creates sub-populations of cells with differential susceptibility to antibiotics. In this thesis, we investigated the effects of functional variability of *M. smegmatis* bacteria on antibiotic tolerance. With the use of live-cell microscopy system that combines imaging, microfluidics, and computational image processing we separated an *M. smegmatis* bulk population into sub-populations with distinct phenotypic characteristics. We then evaluated the response of these sub-populations to rifampicin. We found that there was a correlation between rifampicin tolerance and cell size, cell age, and cell cycle stage. We then conducted PLSR analysis to determine the relative importance of various functional parameters to the rifampicin tolerance. Finally, we created a compartmentalized model framework that is capable of describing the pharmacodynamics of distinct bacterial sub-populations, and could be used to quantitatively resolve single-cell and population-level measurements.

Acknowledgements

I would like to thank my advisor, Dr. Bree Aldridge, for all the guidance and support she gave me during the work on my thesis. I would like to thank Dr. Lauren Black III and Dr. Nikhil Nair for serving on my thesis committee. I would like to thank all the members of the Aldridge Lab, both past and present, especially Owen Bennion, Robert Crutcher, Alex Histed, Michelle Logsdon, and Andrew Min. I would like to thank Jordan Stinson from Dr. David Kaplan's lab, for help with my custom fluidics device. I'd like to thank the Tufts University, in particular the Department of Biomedical Engineering, and the Sackler School of Graduate Biomedical Sciences. Finally, I would like to thank my wife and my parents for all of their support, care, and love.

Table of Contents

List of Tables

List of Figures

The Effects of Functional Variability of *M. smegmatis* on Rifampicin Tolerance

Chapter 1. Introduction

The goal of this project was to investigate various mycobacterial phenotypes, and to determine combined contributions of phenotypic characteristics of mycobacteria to transient insensitivity to antibiotic rifampicin. The work was done by gathering live-cell microscopy data, and analyzing that data with the aid of various software tools (ImageJ, MATLAB, and SIMCA). Furthermore, a compartmentalized model framework was created for describing the pharmacodynamics of distinct bacterial sub-populations.

Chapter 2 provides the background information that explains the biological significance of this study, as well as an overview of bacterial drug tolerance and pharmacodynamic modeling.

Chapter 3 provides a brief review of the current state of knowledge on the functional variability of mycobacteria and its phenotypic-driven drug tolerance.

Chapter 4 describes the main hypothesis that is to be addressed in this thesis and lays out as the various goals of this work.

Chapter 5 describes the materials and methods that were used in this study.

Chapter 6 constitutes the bulk of this thesis. It describes the dataset which was generated and used for analyzing the mycobacterial drug tolerance, and describes the various functional characteristics which were examined: the size and age of the cell during drug treatment start, the size of the cell at birth, the growth rate of the cell, and the cell cycle stages.

Chapter 7 describes the use of partial least-squares regression analysis to investigate the

relationship between the aforementioned functional parameters, and bacterial antibiotic survival outcomes.

Chapter 8 describes the compartmentalized model framework which was created for describing the pharmacodynamics of distinct bacterial sub-population s, and that could be used to quantitatively resolve single-cell and population-level measurements. It also provided a test case which demonstrated the functionality of the model.

Chapter 9 provides the description of a constant flow device that was created for measuring the responses of bulk bacterial populations to antibiotics.

Chapter 10 provides the summary of this thesis.

Chapter 11 describes the future directions of this work, such as further quantifying the contribution of chosen phenotypic parameters to antibiotic tolerance, populating the PD model with single-cell derived data, and creating a reporter for registering the drug-induced death.

Chapter 2. Background

2.1 Clinical Significance

Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis*, which normally affects the lungs. The disease is spread when people sick with pulmonary TB discharge pathogenic microbes into the air, for example, by coughing or sneezing. Overall, a fairly small percentage of people who are infected with *M. tuberculosis* will ever develop active TB. However, once the disease is active, without treatment mortality rates tend to be very high, around 70% of sputum smear-positive cases of pulmonary TB (WHO, 2014).

It is estimated that in the early 1800s, almost all western Europeans were infected with *M. tuberculosis* and approximately one in four deaths were due to TB (Jackson, Mcneil, & Brennan, 2014). Antibiotic TB treatments were first developed in the 1940s, and the most effective first-line antibiotic, rifampicin, was developed by 1960. Antibiotic treatments were hugely successful in drastically reducing TB incidence and mortality rates. Nevertheless, TB remains the world's second deadliest infectious disease, after HIV/AIDS, and still poses a major threat to public health in many parts of the world.

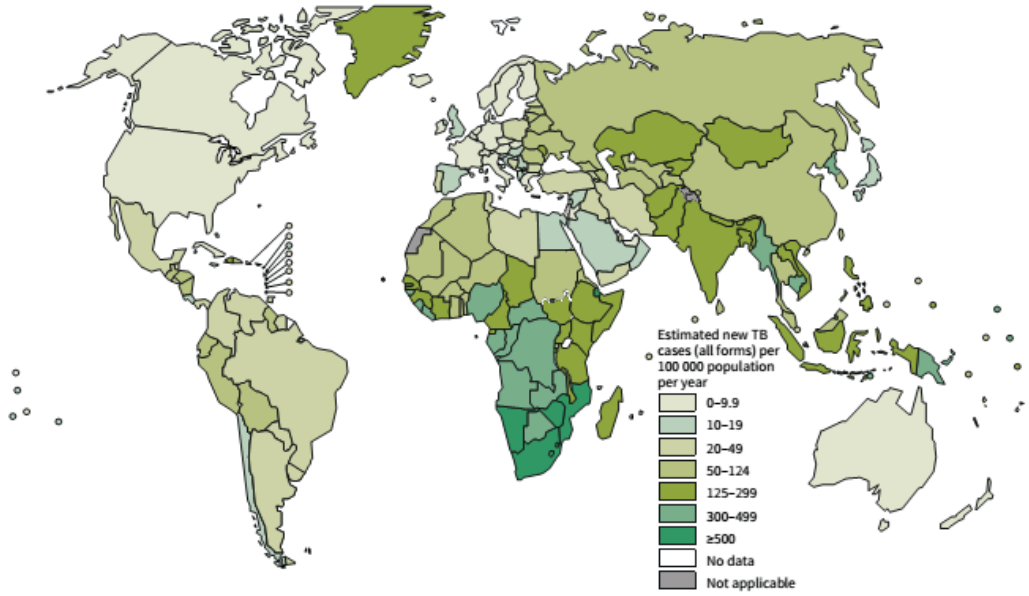


Figure 1: Estimated Incidence of TB Cases Worldwide

TB is most prevalent in Sub-Saharan Africa. Countries in Eastern Europe and Asia also experience high burden of the disease. (WHO, 2014)

In 2013, approximately 9.0 million people developed TB worldwide and 1.5 million died from the disease. An estimated 1.1 million (13%) of the 9 million people who developed TB in 2013 were HIV-positive, of them approximately 360 thousand died (WHO, 2014). The global community is slowly making progress in the fight against the illness: TB incidence rates have slowly been declining in the recent years, and World Health Organization (WHO) estimates that 37 million lives were saved between 2000 and 2013 through effective diagnosis and treatment.

However, drug-resistant TB (DR-TB) poses a new major threat to the control of the disease. In 2013, 3.5% of new and 20.5% of previously treated TB cases were estimated to have had multidrug-resistant TB (MDR TB). In MDR TB the pathogen is resistant to at least isoniazid and rifampicin, the two most potent TB drugs. Current regimens recommended by WHO for treating MDR TB last 20 months, require more expensive and

more toxic therapeutics, and have much lower success rates. At the global level, the MDR TB incidence rate has remained constant in the recent years. However, some areas of the world (especially eastern Europe and the Russian Federation) are experiencing serious MDR-TB epidemics. Additionally, an estimated 9.0% of MDR TB cases had extensively drug resistant TB (XDR-TB), in which in addition to resistance to isoniazid and rifampicin, the pathogen is also resistant to at least one of three injectable second-line drugs (i.e., amikacin, capreomycin, or kanamycin) (WHO, 2014).

Therefore, there exists a clinical need for better understanding of the determinants of antibiotic tolerance. That understanding would allow clinicians to use shorter and more effective treatment regimens, thus leading to improved outcomes.

2.2 TB Drug Resistance

Fundamental challenge in the treatment of TB and in the management of DR-TB is the length of the therapy. For new cases of TB, WHO currently recommends a six-month administration of four first-line drugs: isoniazid, rifampicin, ethambutol, and pyrazinamide. In contrast, the treatment period for bacterium Staphylococcus aureus is on the order of weeks, and for E. coli it is on the order of days (Connolly, Edelstein, & Ramakrishnan, 2007).

The need for multidrug and long-term TB treatment therapy stems from two different drug resistance mechanisms. The first mechanism is the genetic resistance to the drug in the preexisting randomly occurring mutants, a resistance which is heritable and fixed. Simultaneous use of multiple anti-TB drugs makes it less likely that a mutant resistant to a single therapeutic agent will survive the treatment (David, 1970).

The second mechanism of *M. tuberculosis* drug resistance is phenotypic: in patients who relapse early after an appropriate antibiotic therapy, the bacteria remain genetically susceptible to the initial antibiotics and cure is achieved by additional treatment with the same regimen. Since virtually all classes of antibacterial therapeutics require cell replication for their action, the non-replicating bacterial state is thought to render certain sub-population s of genetically homogeneous, antibiotic-susceptible populations of *M. tuberculosis* transiently insensitive to an otherwise effective drug. Therefore, the treatment regiment is required to be long-term, so that all bacteria within the general population eventually leave the antibiotic-resistant phenotypic state (Connolly et al., 2007).

2.3 Pharmacodynamic Modeling

Pharmacodynamics (PD) describes the functional relationship between the drug concentration and some measurable therapeutic effect. In case of antibiotics, that effect is the rate of growth or death of the target bacteria population. PD models of various complexity are an essential tool for drug development and drug usage in the clinical practice.

The simplest PD models describe the effectiveness of a drug treatment with the use of a single parameter, the Minimal Inhibitory Concentration (MIC). It is a point estimate of drug concentration value at which the rate of bacterial population growth is equal to zero. MIC is widely used in the clinical setting, and is the basis for most currently used drug dosing regimens. In the MIC approach, the drugs are normally divided into two broad categories. Antibiotics in the first category, such as vancomycin and macrolides, exhibit mostly time-dependent killing behavior. Their effect is mostly dependant on the amount

of time that the antibiotic is in contact with the bacteria; increasing concentration of the drug past a certain point (generally no more than 4-5 MIC) does not produce an increase in effect. Antibiotics in the second category, such as the amino-glycosides and fluoroquinolones, exhibit mostly concentration-dependant killing behavior. In that case, the bacterial rate of killing increases with an increase in the concentration of the antibiotic, and the effect persists for very high concentrations of the drug (Müller, Dela Peña, & Derendorf, 2004).

However, the MIC is not an optimal marker for developing drug regimens, since it provides very little insight into the kinetics of the therapeutic action. MIC provides information on the change in the number of bacteria at one time point, and because of that, many different combinations of bacterial growth and death rates can have a similar MIC. Furthermore, MIC does not provide adequate information on the bacterial drug resistance, both genetic and phenotype-dependent. Finally, the MIC is essentially an all-or-nothing approach, in which no quantitative distinction is made for various drug concentrations below the MIC.

A far more sophisticated approach to the PD modeling involves the estimation of time-dependent kill curves by exposing target bacteria to a range of antibiotic concentrations, and monitoring the changes in the number of viable cells over time.

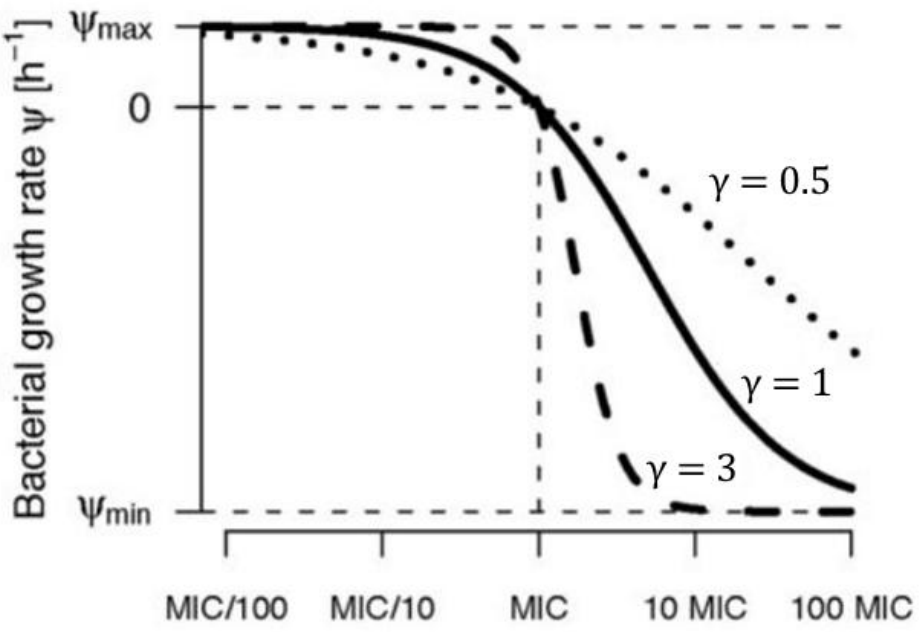


Figure 2: Bacterial Time-Kill Curve

The relationship between the bacterial growth rate and drug concentration. Three curves depict three hypothetical bacterial populations with a varying response function to the change in drug concentration (Regoes et al., 2004)

In this case, the time-dependent relationship between the rate of change in viable bacterial population and drug concentration can be approximated by using mass-action kinetics. The concentration-response relationship is provided by the Hill function:

$$DRUG = \frac{E_{max} C^{\gamma}}{C^{\gamma} + {EC_{50}}^{\gamma}}$$

Where *DRUG* is the antibiotic effect on the growth rate of bacteria (h^{-1}), E_{max} is the maximum drug-mediated death rate (h^{-1}), *C* is the antibiotic concentration (µg/ml), EC_{50} is the drug concentration at which the death rate is at half of its maximum value (µg/ml), and γ is the Hill coefficient, which is an empirical estimation of the sigmoid relationship between the death rate and the drug concentration (Nielsen et al., 2007; Regoes et al., 2004).

This PD model is a far better predictor of the microbiological efficacy of antibiotics than a single PD parameter, such as the MIC, and is a step forward for the development of more robust and effective drug treatment protocols.

Chapter 3. Functional Heterogeneity of Mycobacteria

It has previously been demonstrated that genetically homogeneous mycobacteria colonies exhibit distinct heterogeneous phenotypes. Mycobacteria cells have significantly less symmetrical division than other rod-shaped bacteria. This asymmetry seems to be caused by the fact that mycobacteria cells lack the molecular rulers that ensure symmetric division (Hett & Rubin, 2008).

Furthermore, the available experimental data seem to suggest that two mycobacteria daughter cells that appear after a parent cell division have different elongation rates. Mycobacteria elongate at the cell pole rather than along the lateral axis of the cell, and have one fast growing and one slow growing pole. Therefore, one daughter cell inherits the growing pole, whereas the other cell must synthesize new growth machinery after every division event (Aldridge et al., 2012).

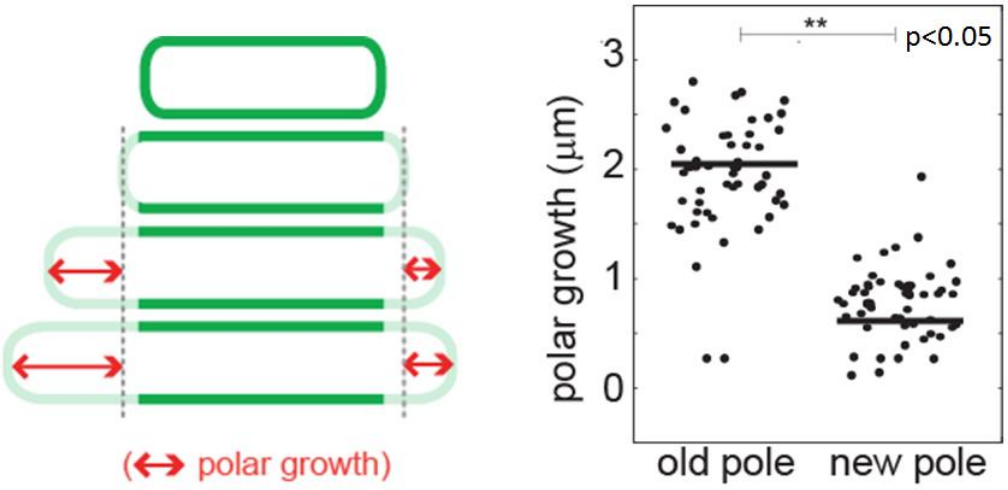


Figure 3: Growth Over One Cell Cycle at New Versus Old Poles

Polar growth rate was assessed using a pulse-chase experiment where cell wall was labeled with amine-reactive dye green dye. The old pole growth rate over one cell cycle is greater than that of the new pole. (From Aldridge et al., 2012. Reprinted with permission from AAAS)

This division pattern creates daughter cells with different growth properties: a larger and faster-growing accelerator (Acc) cell which inherits the old growth pole, and a smaller and slower-growing alternator (Alt) cell which inherits the new pole. Furthermore, successive generations of accelerator cells inherit growth poles of varying ages: some inherit growth poles created in the previous generation, and others inherit growth poles that were created several divisions ago. As the growth pole matures, cells elongate faster and are have a larger birth size (Aldridge et al., 2012).

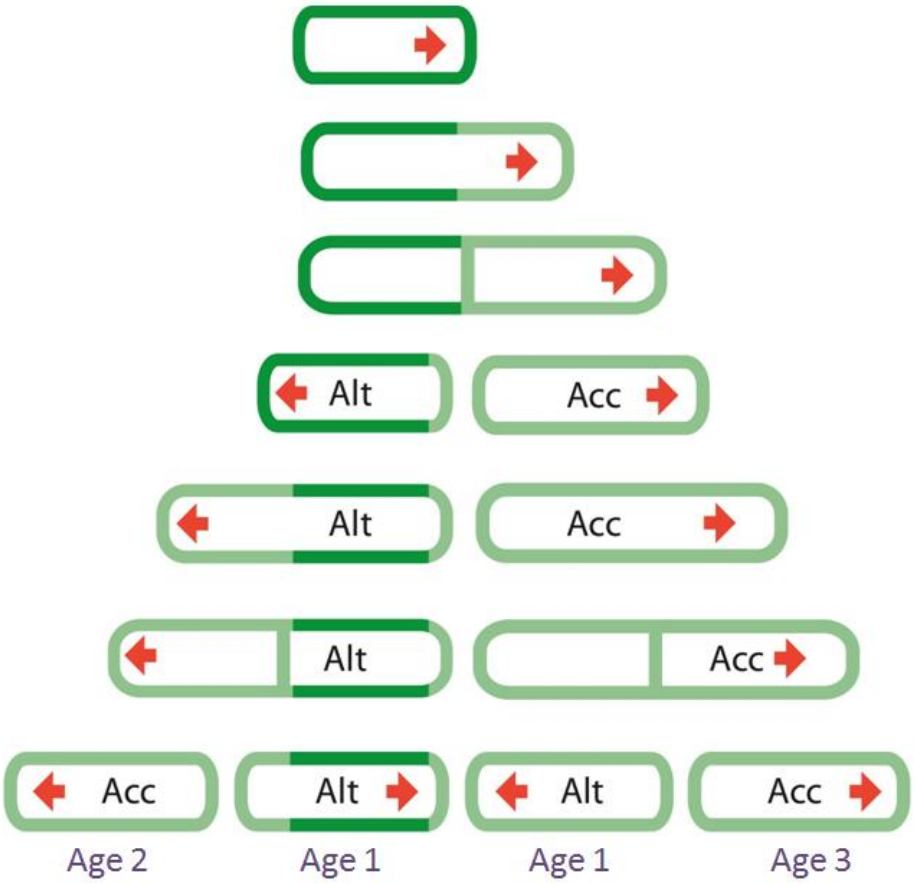


Figure 4: Asymmetric Division Pattern in Mycobacteria

Schematic representation of mycobacterial growth. The elongation rate is much greater at one pole (labeled with a red arrow). Each division event creates two cells: an accelerator (Acc) which inherits the old pole, and an alternator (Alt), which inherits the new pole. The growth pole age is labeled in purple. (Image courtesy of B. Aldridge)

These growth pattern creates rapid functional diversification of closely related cells with variations in elongation rates and sizes. Aldridge et al. showed that this functional variation creates sub-populations of cells with differential susceptibility to antibiotics. The following figure demonstrates the distribution in *M. smegmatis* survival for accelerator and alternator cell microcolonies after the application of meropenem, cycloserine, isoniazid, and rifampicin at their MICs (2.3 mM, 0.04 mg/ml, 25 µM, and 50 µM, respectively). This data seems to suggest that there is a differential antibiotic tolerance for the accelerator and alternator cell phenotypes.

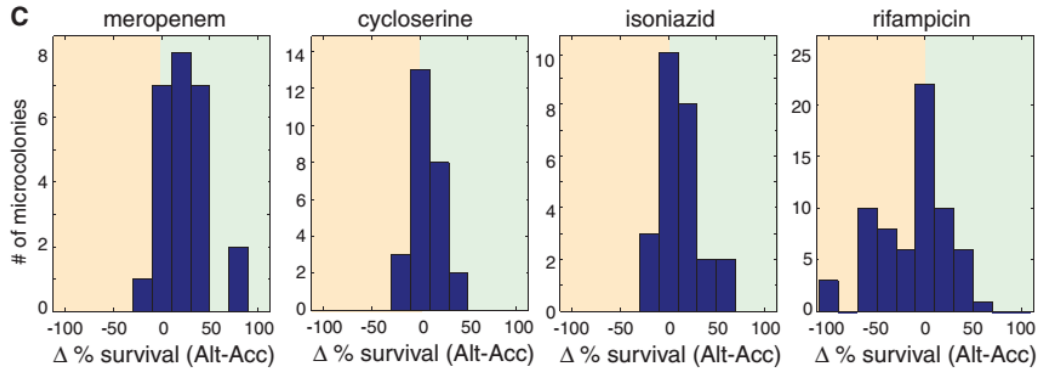


Figure 5: Difference in Bacterial Survival for Acc and Alt Phenotypes

Differential tolerance of antibiotic between Acc and Alt cells. Survival was determining by evaluating the percentage of cells that grew after antibiotic treatment was stopped. (From Aldridge et al., 2012. Reprinted with permission from AAAS)

Furthermore, mycobacterial division cycle seems to be regulated by time rather than by cell size, as in the E. coli. Cell divisions are fairly synchronized in mycobacterial microcolonies, with closely related cells dividing at similar times. The status of cell cycle timing may be assessed with the use of single-stranded binding green fluorescent protein (SSB-GFP), which binds to single-stranded DNA and serves as a replication fork position marker.

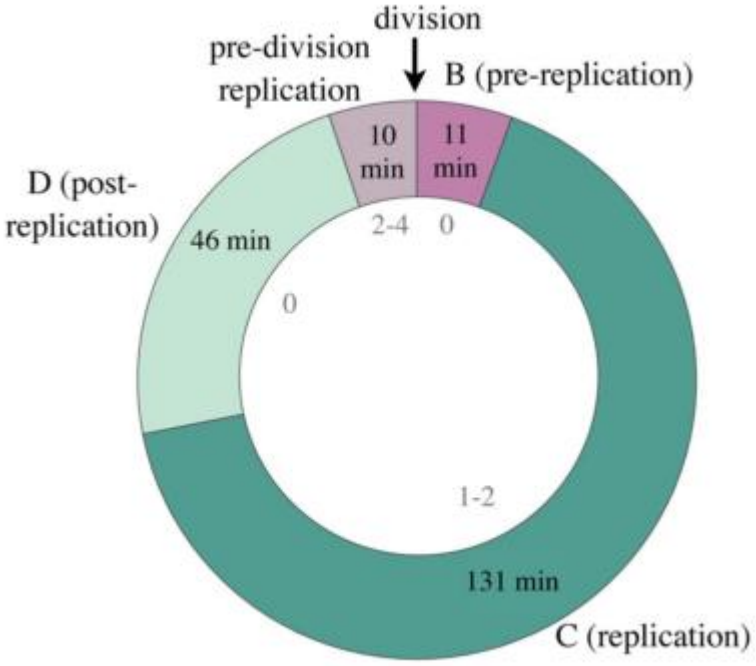


Figure 6: Mycobacterial Cell Cycle

Stages of the mycobacterial cell cycle. The period after the division and before the start of DNA replication is marked as B (G1). The period of DNA replication is marked as C (S). The period between the end of DNA replication and cell division is marked as D (G2). In certain cells DNA replication was again initiated after D, but before division; this stage is labeled as ‘‘pre-division replication." Gray values inside the circle indicate the number of GFP foci, visible in the cell at each cell cycle stage. The average cell cycle stage length is provided. Not all observed cells experienced all cell cycle stages: some lacked B, some lacked D, some lacked E. (Sukumar, Tan, Aldridge, & Russell, 2014)

If certain cell cycle stages are susceptible or tolerant to the drug action, the cell cycle synchronization within individual microcolonies might make most of cells in these microcolonies functionally susceptible or tolerant to drug respectively.

Chapter 4. Designing an Improved Pharmacodynamic Model

Identifying, characterizing, and targeting drug-sensitive and drug-tolerant mycobacteria requires a better understanding of the determinants of antibiotic tolerance. The goal of this project was to investigate phenotypic-dependent antibiotic resistance of mycobacteria, and determine combined contributions of functional characteristics of mycobacteria to transient insensitivity to antibiotic killing.

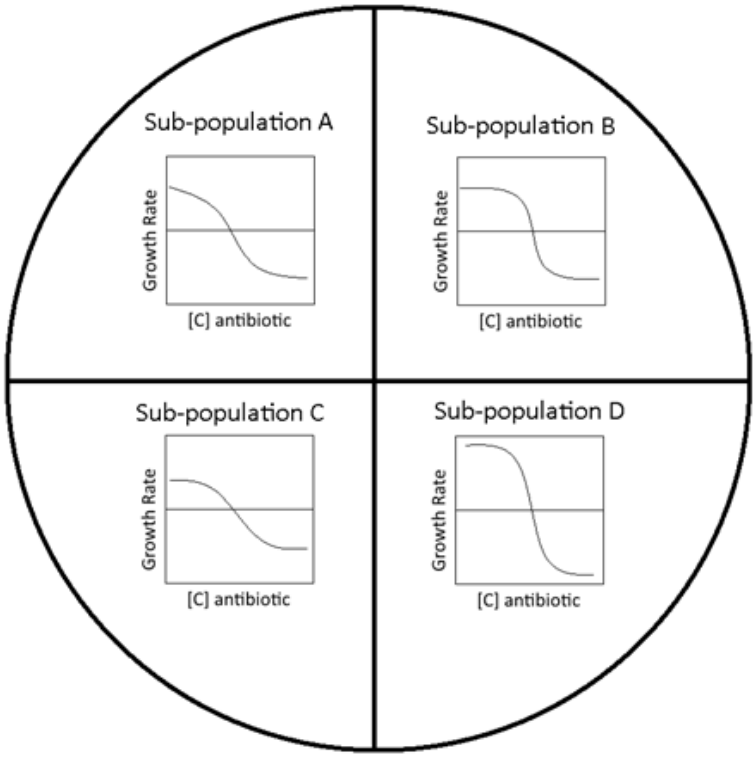


Figure 7: Dividing Bulk Mycobacterial Population into Sub-Populations with Distinct Drug Responses

Let us consider a genetically homogeneous mycobacteria cell population. We hypothesize that based on certain functional parameters, this population can be divided into sub-populations with varying antibiotic kill-curves.

Due to the variable nature of this interaction, we focused on a quantitative, single-cell approach. With the use of live-cell microscopy system that combines imaging, microfluidics, and computational image processing we separated a mycobacteria

population into sub-populations with distinct phenotypic characteristics, and evaluate the antibiotic response of these sub-populations. Furthermore, we created a compartmentalized model framework that is capable of describing the pharmacodynamics of distinct bacterial sub-population s and could be used to quantitatively resolve single-cell and population-level measurements.

The model drug chosen for this project was rifampicin. The mechanism of action of rifampicin is the inhibition of DNA transcription. In the previous work by Adlridge et al. (2012), the plot of functional response to rifampicin had two distinct peaks, which suggests that the phenotypically-derived tolerance to that drug cannot be adequately described by the relatively simple accelerator-alternator model.

Chapter 5. Materials and Methods

5.1 Strains and Cell Culture

Mycobacterium smegmatis was chosen as the model organism for this study. *M. smegmatis* organism is a close nonpathogenic relative of *Mycobacterium tuberculosis*. It is generally accepted that growth and division machinery is conserved between pathogenic and nonpathogenic mycobacteria. The three hour doubling time of *M. smegmatis* is advantageous comparing to the twenty four hour doubling time of *M. tuberculosis*, due to the shortened time required for the experiments.

M. smegmatis strain mc^2155 was transformed with a SSB-GFP replicating plasmid, and augmented with a hygromycin resistance vector, as described previously (Sukumar et al., 2014).

10ml of rich growth medium containing *M. smegmatis* was placed in baffled flasks on a shaker which aerated the culture. Cells were grown at 37°C for 24 hours, allowing them to reach the logarithm growth phase. The culture was then spun down at 2000RPM for six minutes. The resulting cell clump was re-suspended in 200 to 500µl of 7H9 media, depending on the cell pellet size. The cells were then filtered through a 10µm filter to achieve single cell suspension prior to being loaded into the viewing device.

5.2 Medium and Drug Preparation

Rich growth medium (7H9) consisted of 2.35g Difco Middlebrook 7H9 Broth in 450ml of distilled water, which was supplemented with 2ml 50% glycerol, 1.25ml 20% Tween-80, 5ml 10X ADC (albumin, dextrose, catalase), and .5ml Amresco hygromycin B (50 mg/ml).

TCI rifampicin was dissolved in dimethyl sulfoxide to achieve 2mg/ml concentration, and then frozen in 100ml aliquots. Aliquots were stored in a freezer to prevent drug degradation, and were only unfrozen once, immediately prior to the experiment.

To visualize septal membranes, prior to the experiment, an 1.25µg aliquot of Invitrogen FM4-64 FX dye was reconstituted in 400µl of dimethyl sulfoxide (DMSO), and added to 20ml of 7H9 media. Frozen aliquot of TCI Rifampicin (2mg/µl) were thawed and diluted with 7H9 to achieve a pre-determined drug concentration necessary for the experiment.

5.3 Microfluidic Device

A microfluidic viewing device was used to measure the growth and antibiotic susceptibility of mycobacteria at a single-cell level. It allowed cell movement in two dimensions while constraining bacteria to a single focal plane. The polydimethylsiloxane (PDMS) device was cast using molds and then bonded to glass substrates with soft lithography techniques, as described previously (Aldridge et al., 2012; Xia & Whitesides, 1998).

Briefly, the desired pattern was photolithographically defined by using a Mylar mask and used to create masters with a two-layer structure. PDMS prepolymer was mixed with crosslinker at 10:1 weight ratio, stirred, degassed, poured and then cured for a period of 24 hours. Media input and output holes were punched with a 19Ga. flat-tip needle.

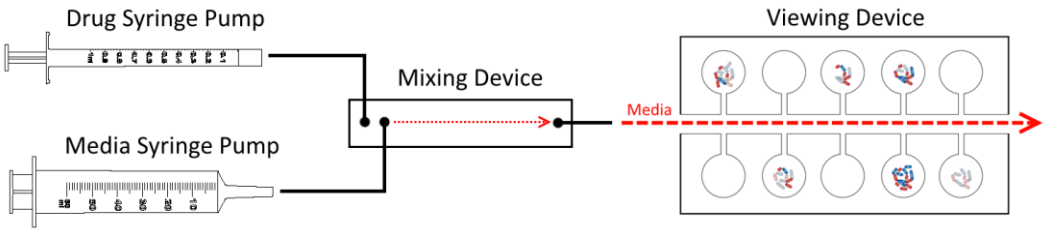


Figure 8: Microfluidics Device

Two syringes (left), controlled by microfluidic pumps, were connected to a mixing device. The media syringe pump ran for the entire duration of the experiment (26 hours); the drug syringe pump was activated ten hours after the start of the experiment, and ran for six hours. The solution from the mixing device flowed through the main channel of the viewing device and into a waste container.

The serpentine mixing device had two inlets, one for 7H9 media, and one for the rifampicin-containing solution. Fluid delivery into the mixing device was accomplished with two microfluidic syringe pumps (Chemyx Fusion 100). The approximate length of the two tubes connecting drug and media syringes to the mixing device was 60cm each. The inner diameter of the tubing was 0.26mm. Both tubes were pre-filled with 7H9 media.

The viewing device contained a main microfluidic media feeding channel, with a height of approximately 10-17µm. Viewing chambers with a diameter of 60 µm were connected to the main channel via side channels 100-200µm long. The height of the side channels and the viewing chambers was approximately 0.8−0.9µm. The heights of these features on the masters mold were determined with a surface profilometer (Dektak ST System Profilometer, Veeco Instruments Inc.).

The solution with suspended mycobacteria cells was loaded into the viewing device with a 1ml syringe. The viewing chambers which contained mycobacteria (ideally just one cell) were selected for the imaging.

5.4 Microscopy

Time-lapse images were acquired using a widefield DeltaVision PersonalDV (Applied Precision, Inc), a modified Olympus IX71 inverted microscope. An automated stage was enclosed in a heated environmental chamber, which was set at a temperature of 37°C. Images were acquired with a 60x- (Plan APO NA 1.42) oil immersion objective. The cells were illuminated with the 461-489 nm InsightSSI Solid State Illumination system (Applied Precision, Inc.) and recorded with a CoolSnap HQ2 camera (Photometric). Images were acquired every fifteen minutes, for a duration of 26 hours. Automated focus correction was performed every five minutes, during imaging time-lapse. Focus was maintained using the hardware-based Ultimate Focus System (Applied Precision, Inc.).

5.5 Antibiotic Treatment

The media syringe delivered 7H9 at the rate of 5µl/min throughout the entire duration of the experiment. *M. smegmatis* were grown in a viewing device for ten hours to establish microcolonies. Cells were then subjected to rifampicin treatment for six hours that achieved a pre-determined drug concentration in the viewing device (refer to Section 6.1 for drug concentrations). When activated, the media syringe delivered the antibiotic solution at the rate of 1µl/min. The minimal inhibitory concentration of rifampicin was determined prior to the experiment by using an alamarBlue assay (Franzblau et al., 1998). Since the tubing was pre-filled with normal growth media, the antibiotic did not reach the viewing device instantaneously. It was empirically determined that from the time of the drug syringe pump activation it took the drug solution approximately half an hour to get to the main channel of the viewing device. Following the drug treatment, the cells were allowed to recover in the normal 7H9 media for ten hours. This technique enabled

imaging of mycobacterial growth for up to five to six generations, at which point the images tended to become too crowded for analysis.

5.6 Image Processing

Images were saved in the Softworx format (Applied Precision, Inc.) and annotated in ImageJ (version 1.47v) with an ObjectJ plug-in (Rasband, n.d.; Vischer, Norbert, Nastase, n.d.). Thirty four microcolonies of *M. smegmatis* (totaling 1020 cells) were analyzed for this study.

For each cell, the cell body length was annotated for each frame, whenever possible. If the cell displayed a visible GFP foci, these were tagged as well, as displayed both in the diagram and live-cell image below:

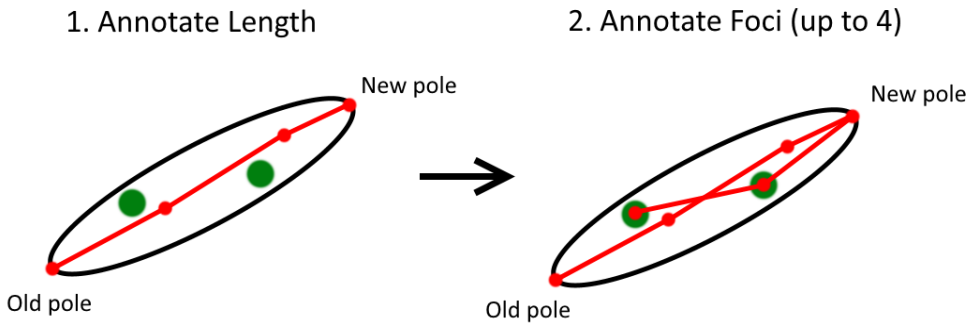


Figure 9: Cell Annotation Workflow

The cell length was marked with a four-point line, from old pole to the new pole. If GFP foci were present, they were annotated with additional points. Up to four GFP foci could be annotated.

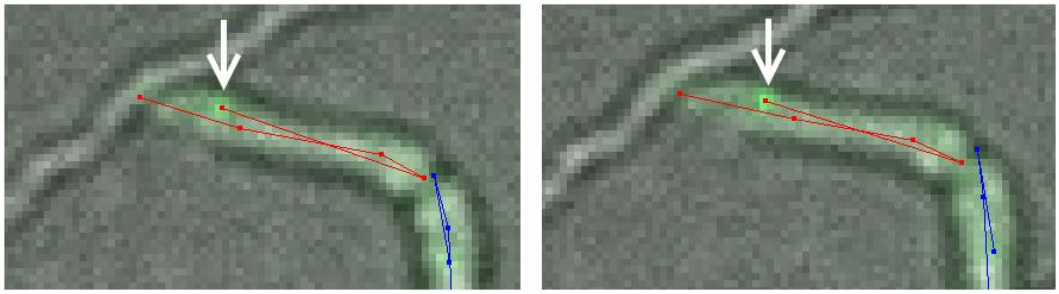


Figure 10: Cell Annotation Example

This brightfield/FITC image shows an *M. smegmatis* cell. The length is annotated by a four-point line (red), a fifth point marks the SSB-GFP focus (marked with a white arrow).

Cell division events were defined as visible "v-snapping" or "pinching" of the cell:

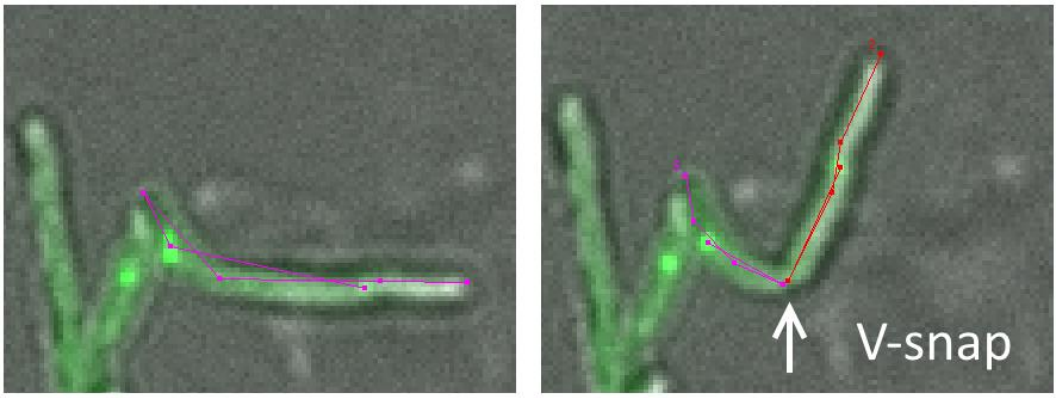


Figure 11: V-snap Division Event

These two images show an annotation transition from a single parent cell (left) to two daughter cells (right), with a V-snap division event.

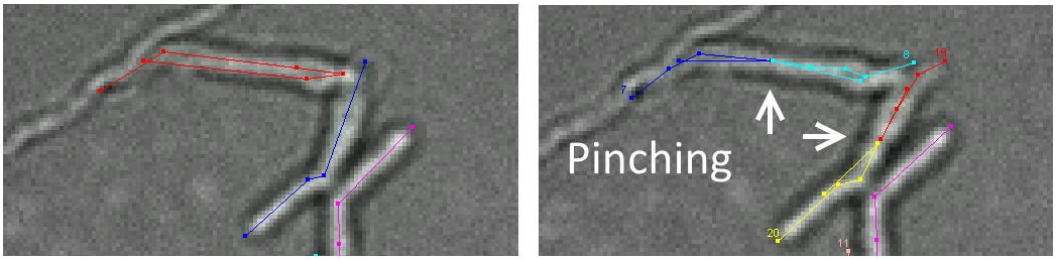


Figure 12: Pinching Division Event

These two images show an annotation transition from a single parent cell (left) to two daughter cells (right), with a pinching division event.

5.7 Data Collation and Analysis

Custom scripts were written in MATLAB (version R2013a) for the purpose of data collation and analysis. The code can be provided upon request. Custom graphic user interface (GUI) was used to assign the following metadata to each cell:

- Microcolony number
- Individual cell number
- Alternator/accelerator status
- Growth pole age (accelerators only)
- Pedigree relationships between cells (i.e. parent/daughter relations)
- Drug treatment outcome

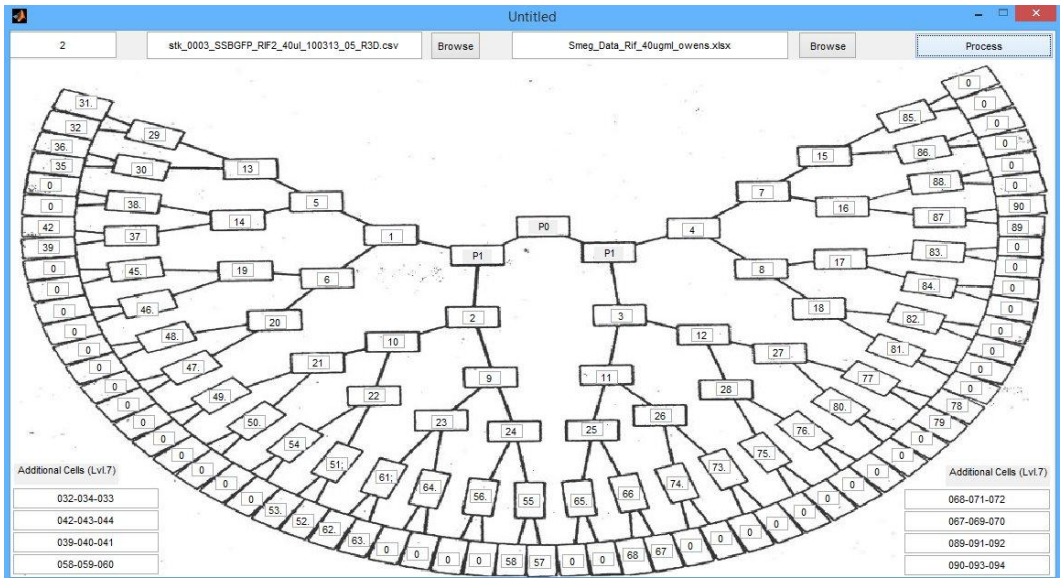


Figure 13: Data Collating GUI

This GUI is based on a paper pedigree chart, courtesy of Owen Bennion.

Drug treatment outcome was defined as follows:

- **Dead cells**[1]. Existed at the time when the drug reached the viewing device, or were born when drug was present in the viewing device. Did not grow or divide after the application of drug until the end of the recording.
- **Live cells.** Existed at the time when the drug reached the viewing device, or were born when drug was present in the viewing device. Grew or produced at least one viable offspring cell before the end of recording.
- **Zombie cells**. Existed at the time when the drug reached the viewing device, or were born when drug was present in the viewing device after the application of drug. Both daughter cells did not grow or divide until the end of the recording.

Please refer to the diagram below for the visual representation of these three sub-populations of cells.

[1] This sub-population of cells does not necessarily include only non-viable cells. It is quite probable that some cells that didn't start growing until the end of the recording were simply dormant.

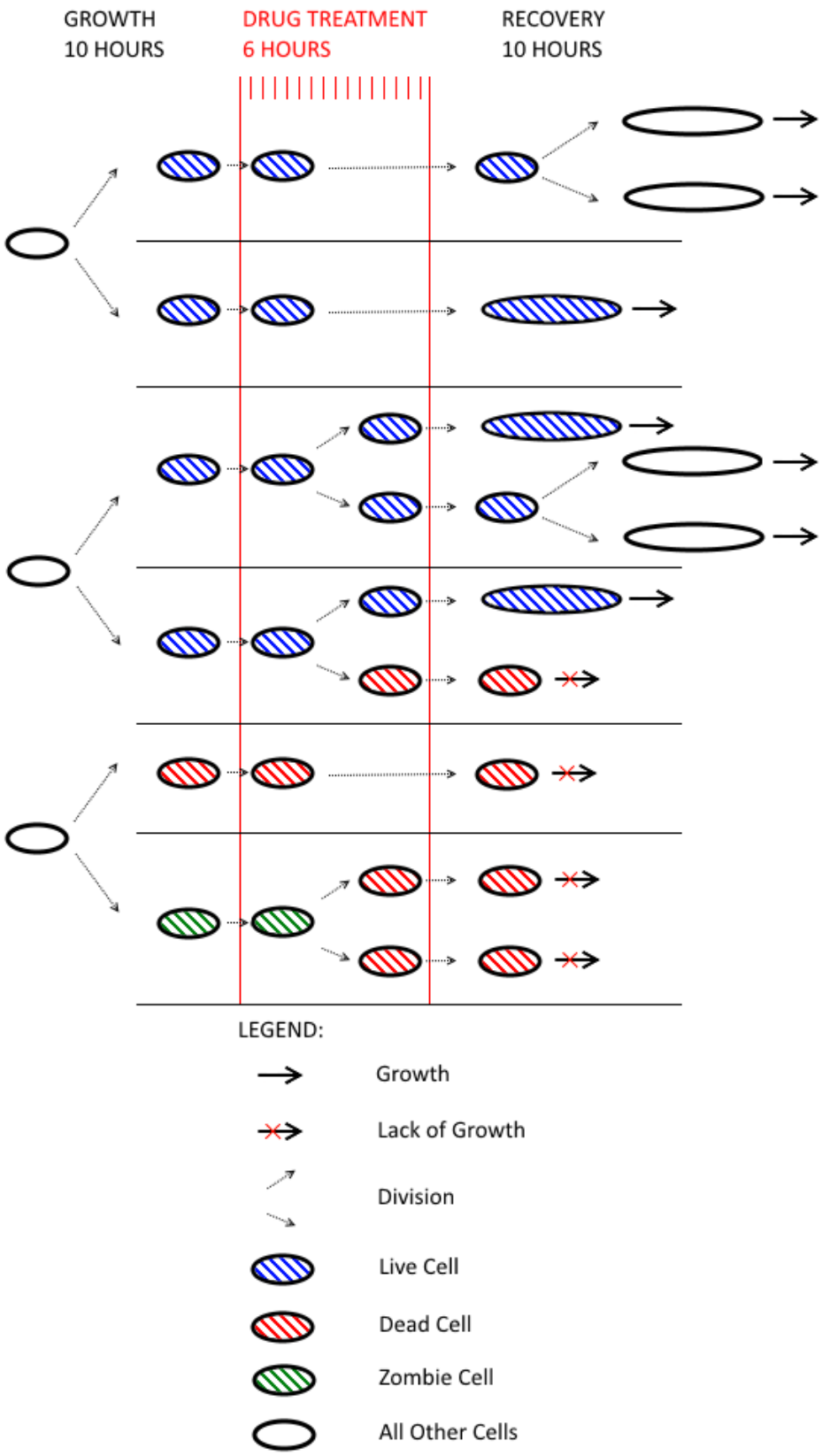


Figure 14: Graphic Representation of Three Drug Treatment Sub-Populations

Besides assigning metadata information to each cell, the script also imported the ImageJ/ObjectJ data, and for each frame where a given cell existed (i.e. between its birth and division or end of recording) it also calculated and assigned the following:

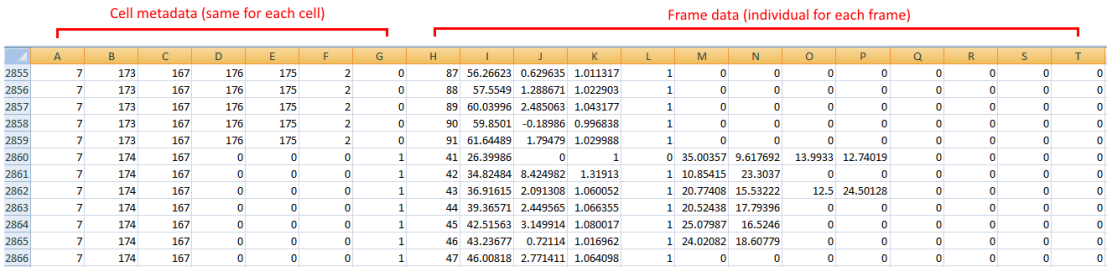
- Cell length (pixels, converted to μm)
- Cell length change from the previous frame (pixels, converted to μm)
- SSB foci presence and position (converted to distance from the cell body poles)

If a certain frame was not annotated for a given cell, the script approximated the length change based on the before- and after- cell length values.

An additional script created graphs with the positions of the SSB foci versus time for each individual cell that had these foci. These graphs were hand-annotated, and the following cell cycle stage data was assigned for each frame:

- B: No SSB-GFP foci after division
- C: 1-2 SSB-GFP foci
- D: period after C with no SSB-GFP foci
- E: 2-4 SSBGFP foci after D and before division

Collated and transformed data was automatically exported into a master file Microsoft Excel spreadsheet; there was a single table for every drug concentration.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
2855	7	173	167	176	175	2	0	87	56.26623	0.629635	1.011317	1	0	0	0	0	0	0	0	0
2856	7	173	167	176	175	2	0	88	57.5549	1.288671	1.022903	1	0	0	0	0	0	0	0	0
2857	7	173	167	176	175	2	0	89	60.03996	2.485063	1.043177	1	0	0	0	0	0	0	0	0
2858	7	173	167	176	175	2	0	90	59.8501	-0.18986	0.996838	1	0	0	0	0	0	0	0	0
2859	7	173	167	176	175	2	0	91	61.64489	1.79479	1.029988	1	0	0	0	0	0	0	0	0
2860	7	174	167	0	0	0	1	41	26.39986	0	1	0	35.00357	9.617692	13.9933	12.74019	0	0	0	0
2861	7	174	167	0	0	0	1	42	34.82484	8.424982	1.31913	1	10.85415	23.3037	0	0	0	0	0	0
2862	7	174	167	0	0	0	1	43	36.91615	2.091308	1.060052	1	20.77408	15.53222	12.5	24.50128	0	0	0	0
2863	7	174	167	0	0	0	1	44	39.36571	2.449565	1.066355	1	20.52438	17.79396	0	0	0	0	0	0
2864	7	174	167	0	0	0	1	45	42.51563	3.149914	1.080017	1	25.07987	16.5246	0	0	0	0	0	0
2865	7	174	167	0	0	0	1	46	43.23677	0.72114	1.016962	1	24.02082	18.60779	0	0	0	0	0	0
2866	7	174	167	0	0	0	1	47	46.00818	2.771411	1.064098	1	0	0	0	0	0	0	0	0

Figure 15: Master File Layout

Columns A-G will repeated for every single individual cell, and contained cell's metadata:

- A - microcolony identifier.
- B - cell number
- C - parent cell number
- D, E - daughter cell numbers
- F - pole age. 0 is alternator, 1 is accelerator with a pole age of 1, 2 is accelerator with a pole age of 2, etc.
- G - drug treatment outcome. 0 is live, 1 is dead, 2 is zombie.

Columns H-P contained data for each annotated frame:

- H - frame number
- I - length of the cell at the given frame number (in pixels)
- J - absolute length change of the cell from the previous annotated frame number (in pixels).
- K - relative length change of the cell from the previous annotated frame number (i.e. current length in pixels divided by the previous length in pixels).
- L - The difference between the present frame and the previously annotated frame.
- M-T - foci distance to new and old poles.

Additional MATLAB scripts were used to further analyze this data.

PLSR analysis was performed with a demo-version of SIMCA (version 14).

Chapter 6. Results of Data Analysis

6.1 Dataset Size

The following table provides the experimental sample size which was used for data analysis:

Rifampicin Concentration (µg/ml)	xMIC	Microcolonies (n)	Cells (n)
20	0.5	6	138
40	1	7	270
80	2	7	216
120	3	14	396

Table 1: Total Sample Size

6.2 Characterizing *M. smegmatis* in the Absence of Drug

Data analysis was conducted for all cells that existed strictly prior to the application of rifampicin, as a method of control. If the cell existed after ten hours since the beginning of the recording (thus being subjected to drug action), it was not included into this data set.

The majority of *M. smegmatis* cells divided about 2:15 - 3:30 hours after the cell birth. The average time of division was just under three hours.

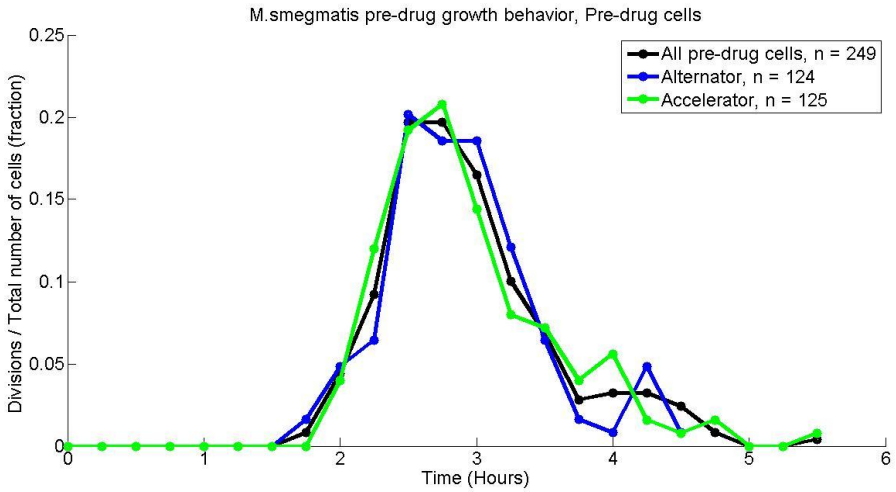


Figure 16: *M. smegmatis* Division Times (No Drug)

The following table provides data on the average birth and division sizes of *M. smegmatis* cells. As expected, the accelerator cells tended to be larger at birth and larger still at the division when compared to the alternator cells.

	All Cells	Accelerators	Alternators
Mean Birth Size (µm)	3.5±0.77	3.9±0.73	3.2±0.61
Mean Division Size (µm)	6.3±1.25	6.9±1.19	5.7±0.97
Sample Size (n)	249	125	124

Table 2: *M. smegmatis* Birth and Division Sizes

The following figure provides time-corresponding data for an average *M. smegmatis* cell from time of birth. The recorded data for all available cells was "stacked" and began at 0:00 hours, i.e. the data below represents an average size of the cell from the time of birth. The drop in size after four hours can be contributed to the fact that most cell have undergone division by that point (refer to Fig. 15), and the remaining "lagging" cells skewed the average size downward.

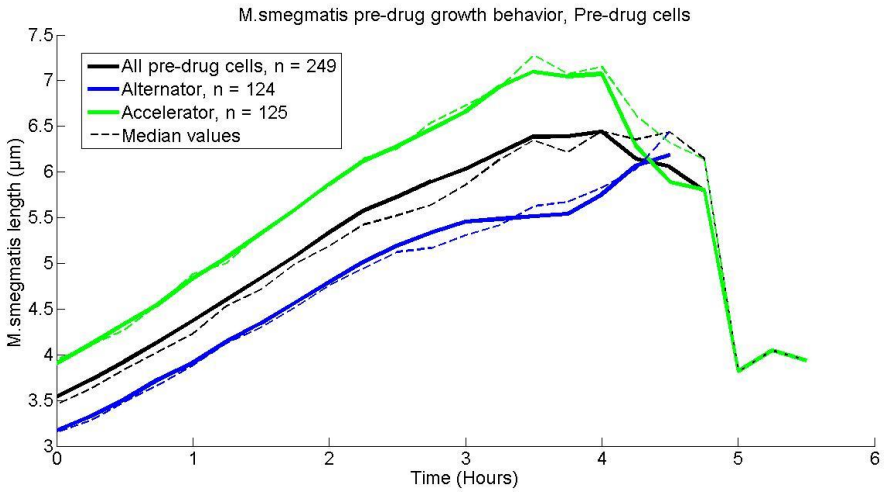


Figure 17: Birth and Division Sizes (No Drug)

The following figure provides the average growth rate for an *M. smegmatis* cell, i.e. average rate of cell elongation at 15 minute intervals. The accelerator cells tended to grow at a faster rate when compared to the alternator cells. Additionally, a polynomial curve with one degree of freedom (i.e. a linear plot) was fitted for the period of 0:00-3:00 hours, in order to capture data for most cells within the sample. The average rate of growth increased as the cells aged.

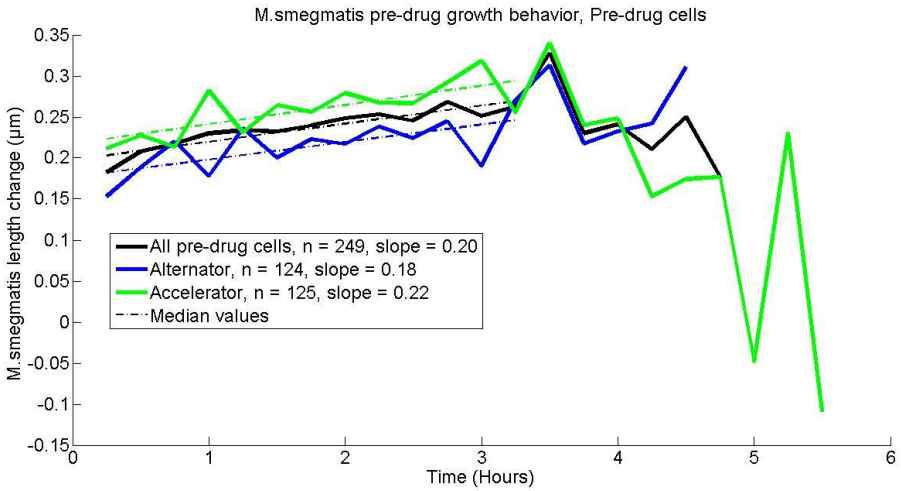


Figure 18: Growth Rate (No Drug)

The following figure provides the average length of accelerator cells with various growth pole ages. The average size of the accelerator cell seemed to be greater for older pole ages.

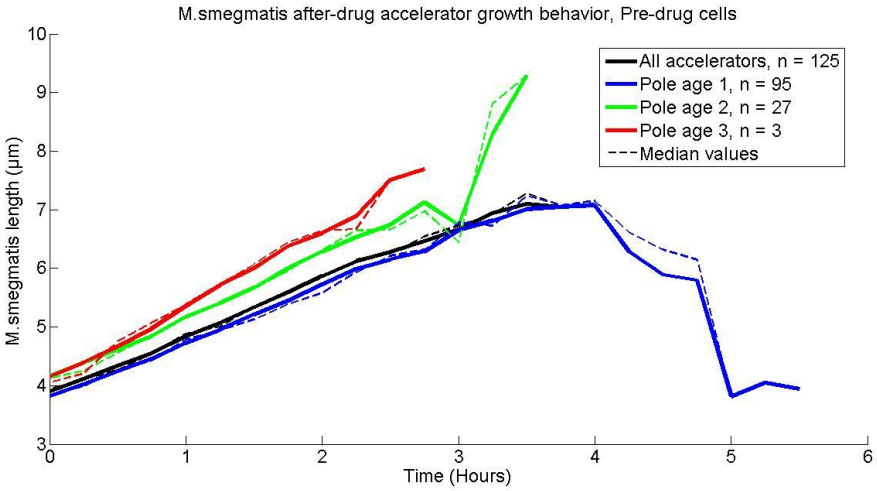


Figure 19: Average Size for Various Accelerator Growth Pole Ages (No Drug)

The following figure provides the average growth rate of accelerator cells with pole age of one, and a combined population of accelerator cells with pole ages of two and three. The accelerators with older pole ages tended to growth faster.

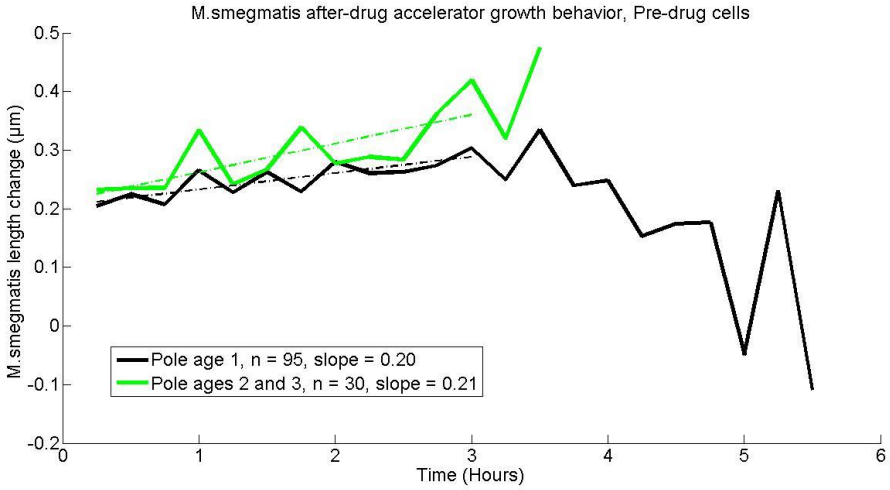


Figure 20: Accelerator Average Growth Rate for Pole Ages (No Drug)

6.3 Accelerator and Alternator Cells Exhibit Variable Tolerance of Rifampicin

The following data analysis examines the relationship between the accelerator and alternator phenotypes, and tolerance of rifampicin at four different concentrations of the drug.

Rif [C] (µg/ml)	**Acc** Cells (n)	Live (n)	Dead (n)	Zombie (n)	% Tolerant
20	42	18	19	5	42.9%
40	67	20	40	7	29.9%
80	51	26	19	6	51.0%
120	127	29	77	21	22.8%

Table 3: Drug Tolerance in Accelerator Cells

Rif [C] (µg/ml)	**Alt** Cells (n)	Live (n)	Dead (n)	Zombie (n)	% Tolerant
20	40	14	20	6	35.0%
40	68	5	52	11	7.4%
80	51	13	31	7	25.5%
120	120	9	95	16	7.5%

Table 4: Drug Tolerance in Alternator Cells

At each drug concentration, a majority of the live cells sub-population are accelerators. This difference is statistically significant[2] at $p<0.05$ for three drug concentrations (40, 80, and 120 µg/ml), with p-values of 0.0027, 0.0374, and 0.0012 respectively.

The following figure examines what percentage of tolerant cells were accelerators and alternators, and what percentage of susceptible cells were accelerators and alternators at the drug concentration of 120 µg/ml. Accelerators tended to be much more tolerant of rifampicin.

[2] Due to the bi-modal state of live cells (i.e. accelerator vs. alternator), a "fair-coin" χ^2 test was used, with a null hypothesis that a live cell has an equal probability of being either an accelerator or an alternator.

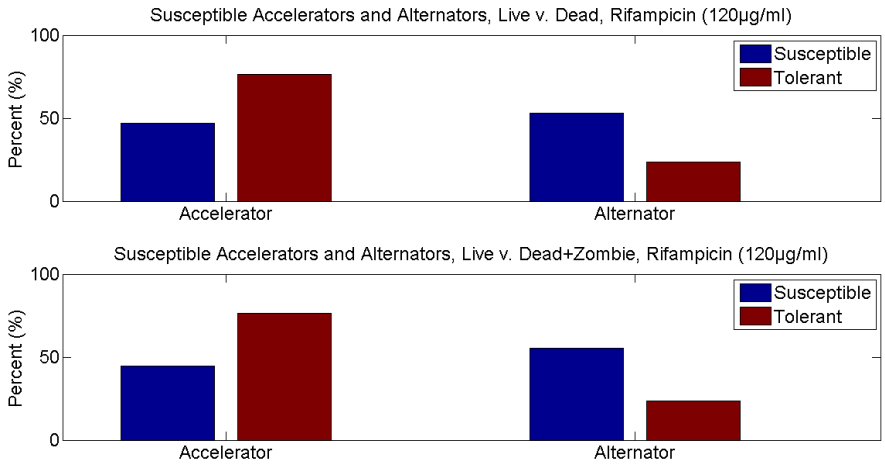


Figure 21: Accelerator and Alternator Drug Tolerance

The susceptible population pictured in the top plot includes only the cells that were defined as dead. The susceptible population pictured in the bottom plot includes both the cells that were defined as dead, and cells that were defined as zombies. The second plot was created to see, whether expanding the definition of susceptible cells to both dead and zombie sub-populations would alter the accelerator to alternator ratio. No discernible difference can be seen between the two definitions of susceptible populations..

6.4 Cell Size at Drug Treatment Start Strongly Correlates with Tolerance

The following data visualization depicts the averaged lengths of six bacterial sub-populations at each recorded time point. Judging from the graph, at the time of drug treatment start (t=10.5 hours), the average size of the dead cell sub-population was smaller than that of the live cell sub-population. The zombie cell sub-population was on average larger than both dead and live cells.

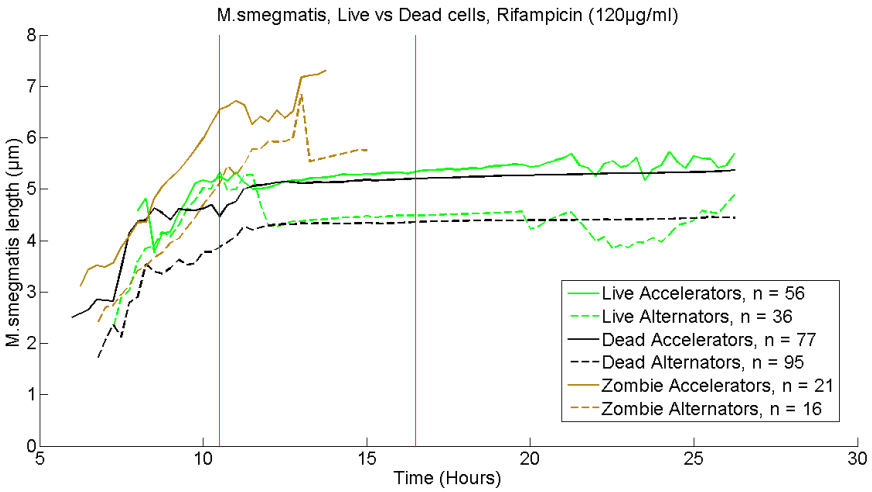


Figure 22: Sub-Population Average Size at Each Frame[3]

The sub-populations were grouped by outcome (live/dead/zombie), as discussed in Section 5.7. Additionally, these three sub-populations were further divided into accelerators and alternators. The start and end time of drug treatment (120µg/ml) is designated by the vertical red lines.

Next figure shows the box plot of the average sizes of aforementioned sub-populations at the time of drug treatment start, for cells that were treated with rifampicin at 120µg/ml. There is a statistically significant difference between the mean size of live and dead cells at the time of drug treatment start. Wilcoxon signed rank sum was used instead of a paired samples t-test due to a relatively small sample size for the live cell cohort.

[3] This figure includes data for cells that were born after the drug treatment ended. These cells are grouped together with live cell sub-population in order to make the graph easier to read.

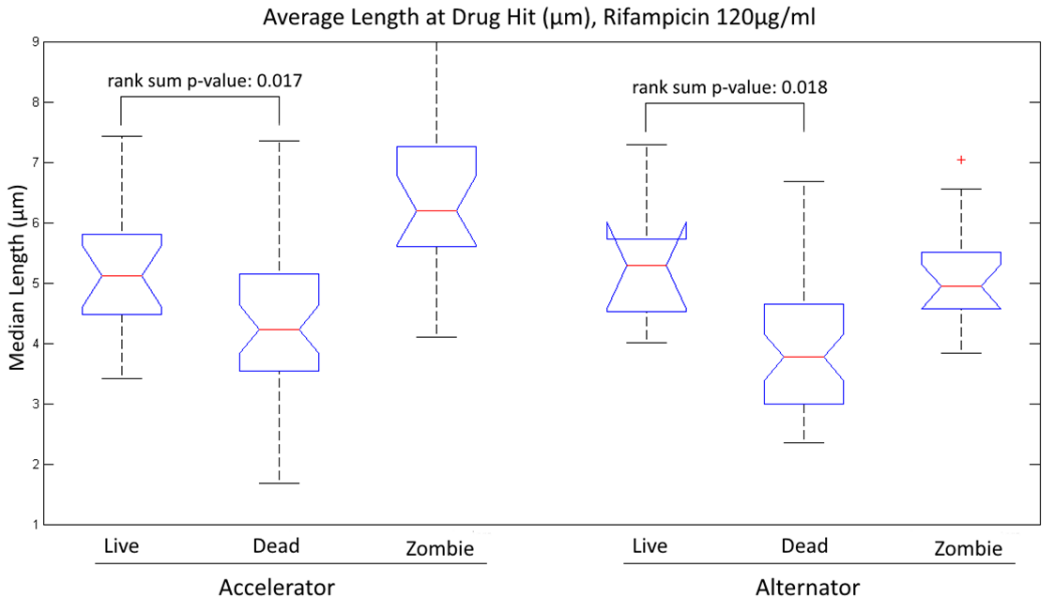


Figure 23: Sub-Population Mean Size at Rifampicin Treatment Start[4]

For rifampicin concentration of 120µg/ml. Central red lines show medians for the given sub-population. The edges of the box are the 25th and 75th percentile values. The whiskers extend to the most extreme values not considered outliers by the MATLAB *boxplot* algorithm. The outliers are plotted individually, as red crosses. The notches signify the 95% significance level (based on a normal distribution assumption).

A similar relationship between dead, live, and zombie cell sizes can be observed at smaller drug concentrations. However, the difference between live and dead sub-populations is no longer statistically significant at p=0.05 for a Wilcoxon signed rank sum test.

[4] For this graph, sub-populations of live, dead, and zombie cells had to be limited only to cells that existed at the start of drug treatment.

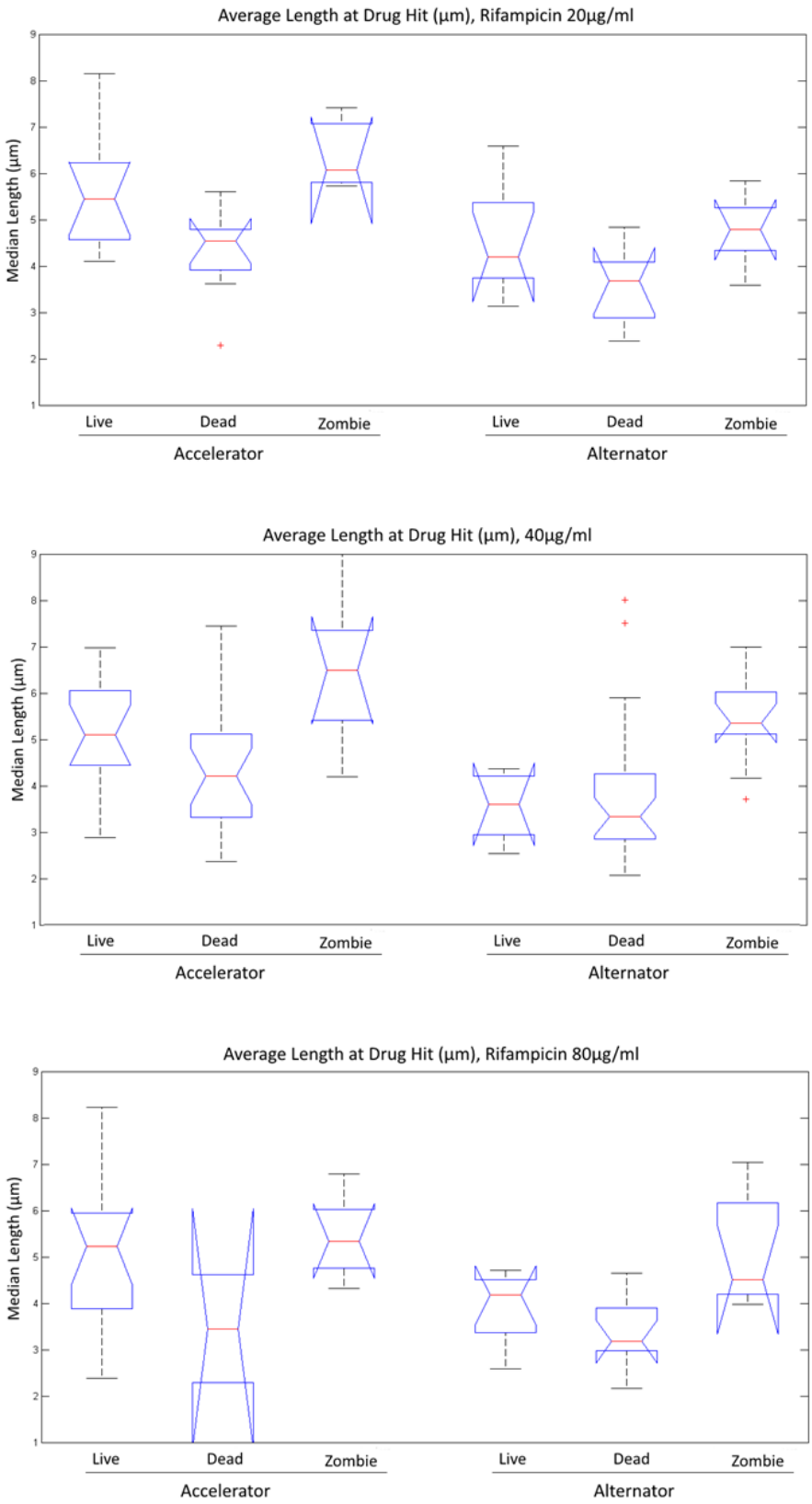


Figure 24: Box Plot of Sub-Population Mean Sizes at Rifampicin Treatment Start

The following graph provides the additional information on the sizes of various cell sub-populations at the time of the drug treatment start. It is a histogram distribution of cell sizes for six aforementioned sub-populations, along with two control histograms (all accelerator and all alternator cells sizes 1.5 hours before the beginning of drug treatment).

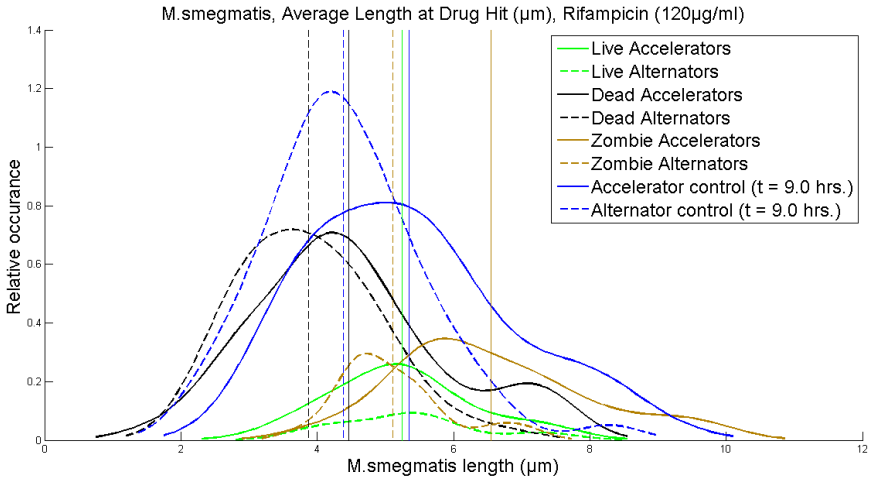


Figure 25: Histogram of Sub-Population Mean Size During Treatment Start

The histogram fit was created with the MATLAB *histfit* function (one hundred bins, kernel smoothing function fit). Y-axis depicts the relative occurrence of a certain bacterium size value at the beginning of drug treatment, i.e. the number of cells sizes at that particular bin. The vertical lines represent the mean sizes for each sub-population.

The data seems to suggest that *M. smegmatis* size at the start of drug treatment correlates with tolerance outcome. Cells that tended to not grow during the recovery period were on average small at the time of antibiotic treatment start; cells which had non-viable offspring were on average large. The average size of cells that grew during the recovery period was somewhere between the two aforementioned extremes.

There could be multiple underlying functional factors that lead to this size difference, for example: different growth rate, different birth size, or different birth timing. The next

three sections will examine the possible connection of these three factors to drug tolerance.

6.5 No Apparent Relationship Between Growth Rate and Drug Tolerance

The following data visualization depicts the averaged growth rates of six bacterial sub-populations at each available time point.

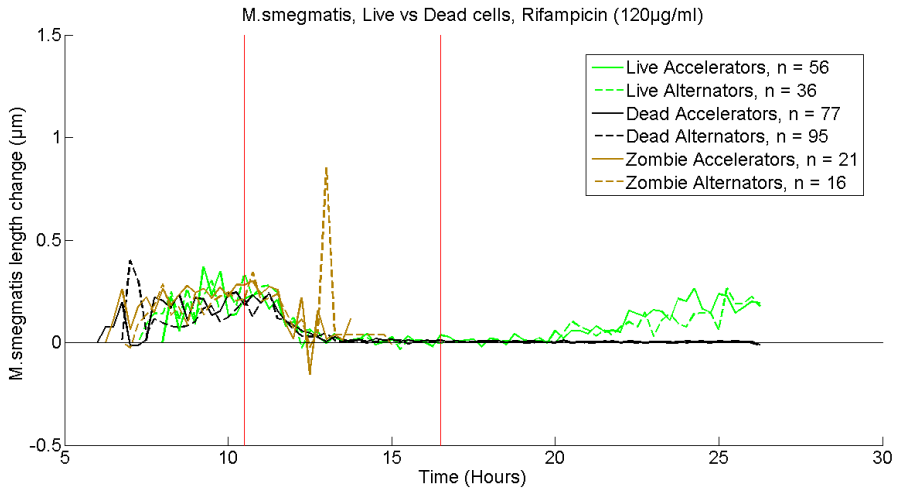


Figure 26: Average Sub-Population Growth at Each Frame[5]

The sub-populations were grouped by outcome (live/dead/zombie), as discussed in Section 5.7. Additionally, these three sub-populations were further divided into accelerators and alternators. The start and end time of drug treatment (120µg/ml) is designated by the vertical red lines.

There is no discernible difference in the growth rates of the six sub-populations of *M. smegmatis* cells. It is interesting to note, that live cells started growing again about 3-4 hours after the drug syringe was stopped.

6.6 Cell Birth Size Weakly Correlates with Tolerance

The graph below depicts individual cell data for the aforementioned six sub-populations, for cells that were treated with rifampicin at 120µg/ml.

[5] This figure includes data for cells that were born after the drug treatment ended. These cells are grouped together with live cell sub-population in order to make the graph easier to read.

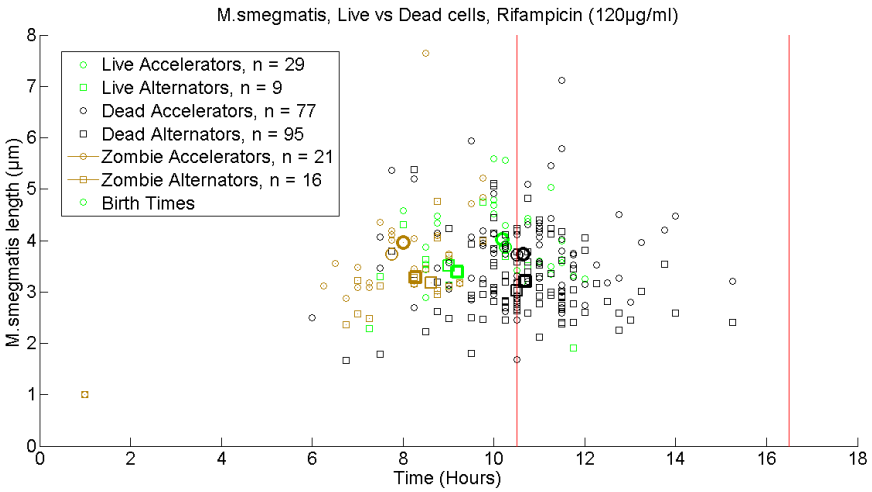


Figure 27: Individual Cell Birth Times and Birth Sizes

The Y-axis corresponds to the birth size of individual bacteria. Bolded markers provide averages for each of the six sub-populations; thicker bold marker in the mean, and thinner bold marker in the median of the corresponding sub-population. The start and end time of drug treatment (120µg/ml) is designated by the vertical red lines.

Next figure shows the box plot of average birth times of the aforementioned sub-populations.

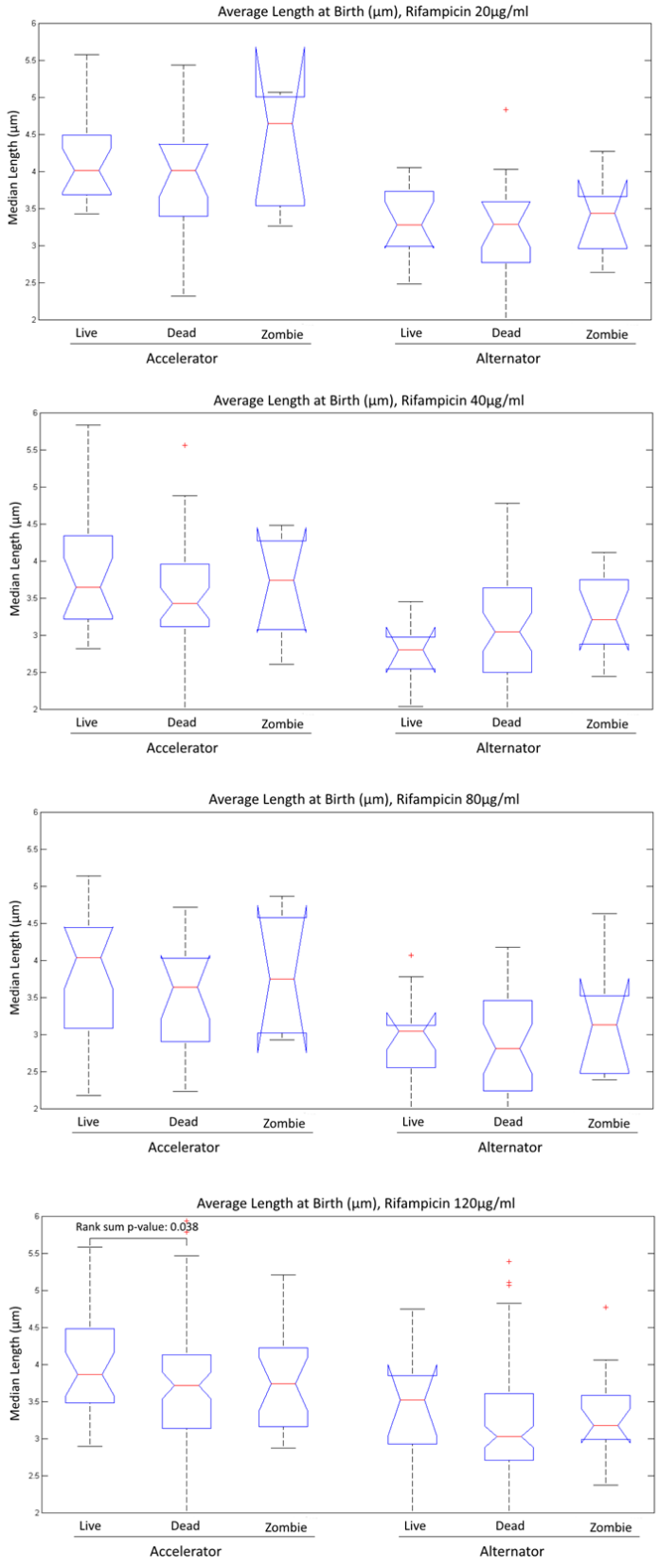


Figure 28: Sub-Population Mean Sizes at Birth

There is no strong pattern of size at birth differences between the various functional sub-populations, although cells that ended up not growing during the drug recovery time did tend to have smaller birth size. In one case, the rank-sum test did show statistical significance at $p<0.05$ between the birth size of live accelerators and dead accelerators, but the p-value was a relatively modest 0.038.

The following graph provides the histogram distribution of birth sizes for various sub-populations, along with two control histograms, birth sizes of all accelerator and alternator cells that existed only between 4 and 10 hours since the start of the experiment.

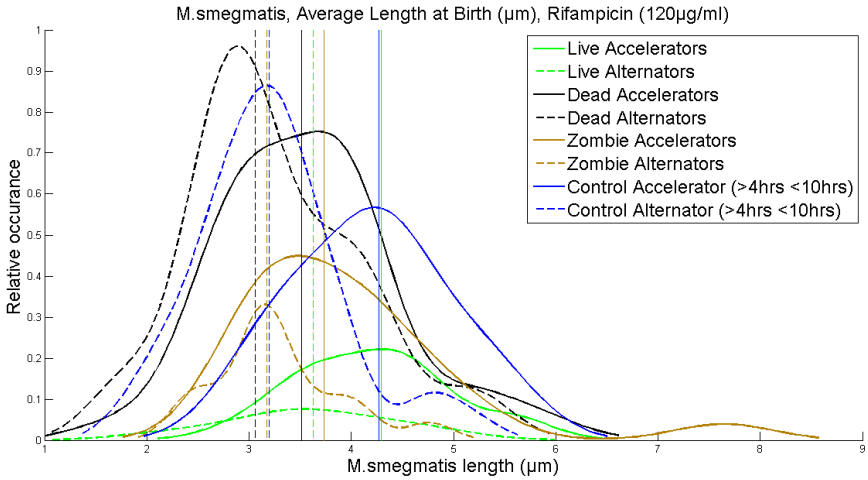


Figure 29: Histogram of Sub-Population Mean Sizes at Birth[6]

The histogram fit was created with the MATLAB *histfit* function (one hundred bins, kernel smoothing function fit). Y-axis depicts the relative occurrence of a certain bacterium size value at the beginning of drug treatment, i.e. the number of cells sizes at that particular bin. The vertical lines represent the mean sizes for each sub-population.

6.7 Cell Birth Time Strongly Correlates with Tolerance

The graph below depicts individual cell data for the aforementioned six sub-populations, for cells that were treated with 120µg/ml Rifampicin.

[6] The histograms do not include cells that were born after the beginning of drug treatment, since the drug stress affected their birth size, thus skewing the birth size distributions.

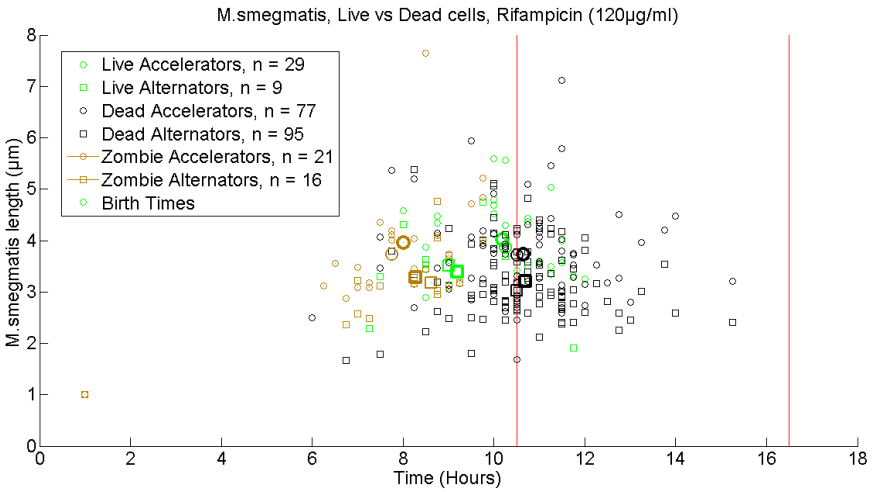


Figure 30: Individual Cell Birth Times and Birth Sizes

The X-axis corresponds to the birth time of individual bacteria. Bolded markers provide averages for each of the six sub-populations; thicker bold marker in the mean, and thinner bold marker in the median of the corresponding sub-population. The start and end time of drug treatment (120µg/ml) is designated by the vertical red lines.

Next figure shows the box plot of average birth times of the aforementioned sub-populations, relative to the drug treatment start. The dead cells tended to be born a short time before the drug treatment start; zombie cells tended to be born a long time before drug treatment start; live cells' birth times fell somewhere in-between the dead and zombie birth times. The rank-sum test shows significant difference between live and dead accelerator cells for 80µg/ml treatment ($p<0.05$), and significant difference between live and dead alternator cells for 120µg/ml treatment ($p<0.05$).

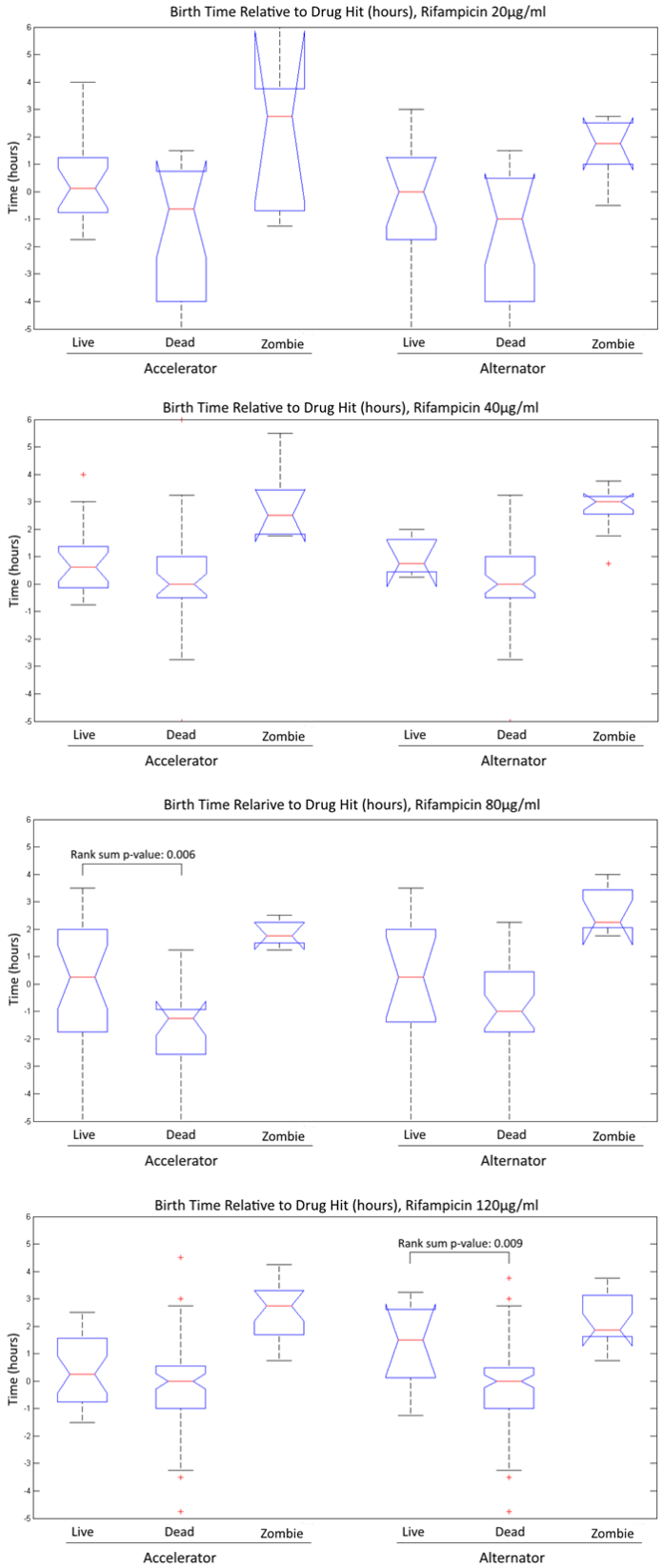


Figure 31: Sub-Population Mean Birth Times Relative to Treatment Start

The following graph provides the histogram distribution of birth times for various sub-populations, relative to the drug treatment start. There is no independent control; however there is a histogram of all combined birth times for accelerators and alternators prior to their division into dead, live, and zombie populations.

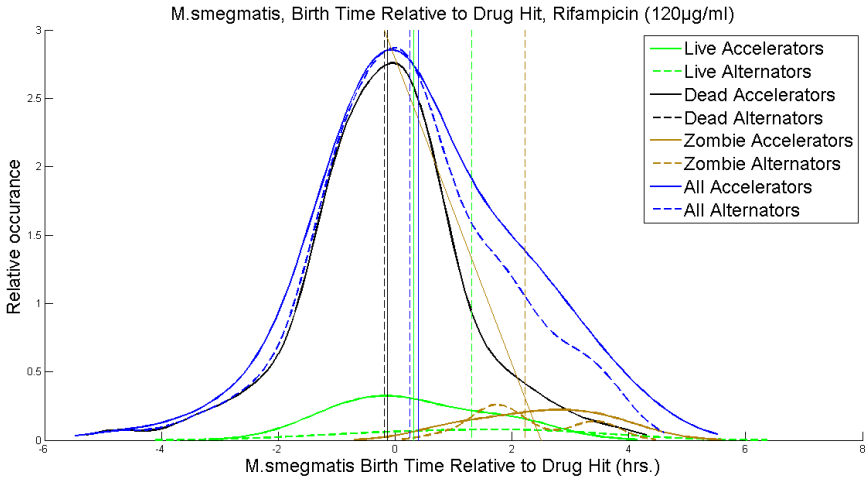


Figure 32: Histogram of Sub-Population Mean Birth Times Relative to Drug Treatment Start

The histogram fit was created with the MATLAB *histfit* function (one hundred bins, kernel smoothing function fit). Y-axis depicts the relative occurrence of a certain bacterium size value at the beginning of drug treatment, i.e. the number of cells sizes at that particular bin. The vertical lines represent the mean sizes for each sub-population.

6.8 Cell Cycle Stage Strongly Correlates with Tolerance

Since cell age during the start of drug treatment seemed to be an important factor, and since the cell cycle in mycobacteria seems to be time-dependent, we decided to investigate whether the stage of cell cycle at the start of drug treatment would correlate with rifampicin tolerance.

The following plot depicts the cell cycle stage for three *M. smegmatis* sub-populations (live, dead, zombies), for 120µg/ml drug treatment. Cells that did not display any foci at

any time were not included, since it is impossible to determine what stage of cell cycle they were in at t = 10.5 hours.

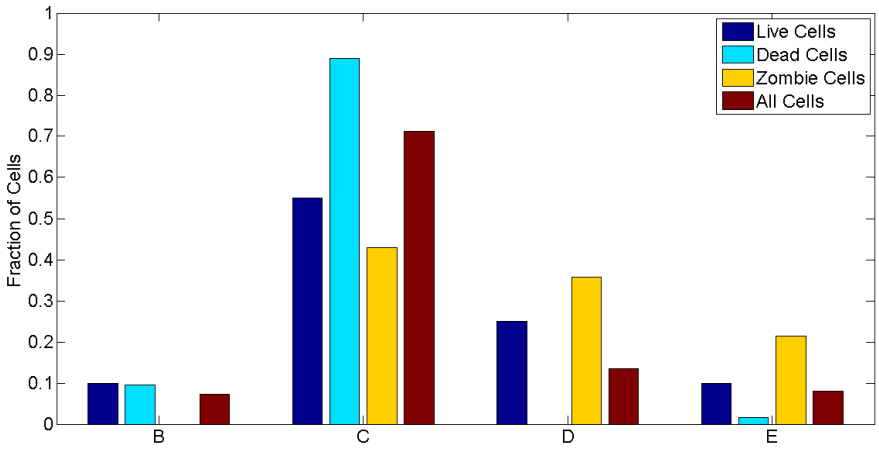


Figure 33: Sub-Population Cell Cycle at the Time of Drug Treatment[7]

The dead cell sub-population (n=91) tended to be in the early stages of the cell cycle, while the zombie cell sub-population (n=29) tended to be in the late stages of the cell cycle. The live cell population (n=30) tended to have a distribution close to the distribution of a normal cell cycle. The "all cell" population is the cell cycle distribution of all the cells that were present at the time of drug treatment start.

The following plot depicts the C cycle stage for three *M. smegmatis* sub-populations (live, dead, zombies), at the time of 120µg/ml rifampicin drug treatment start. Cells that did not display any foci were not included, since it is impossible to determine what stage of cell cycle they were in at t=10.5 hours.

[7] The graph depicts the cells that existed at the time of the drug treatment start.

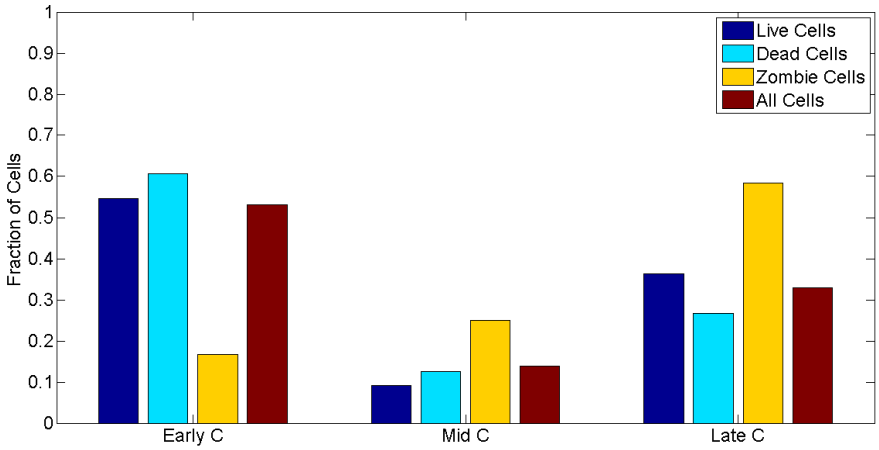


Figure 34: Sub-Population C Cell Cycle at the Time of Drug Treatment

The graph depicts the cells that were in C cell cycle. Early C was defined as 0-30 minutes, mid C was defined as 45-75 minutes, late C was defined as anything over 90 minutes (the images were taken every 15 minutes). Sample size: n=11 for live cells, n=56 for dead cells, n=12 for zombie cells. The zombie cell sub-population tended to be in the later stages of the C cell cycle.

Chapter 7. Analyzing the Impact of Various Mycobacterial Functional Characteristics on Drug Tolerance with Partial Least-Squares Regression

The data in the previous chapter suggests that there is a correlation between functional *M. smegmatis* differences, and rifampicin tolerance. In particular, size of the cell and the stage of cell cycle at the time of drug treatment start seems to be a good predictor of rifampicin tolerance. Birth timing seems to correlate with drug tolerance as well, though to a slightly lesser degree. Naturally, these functional parameters aren't independent of one another, e.g. if the cell is born shortly before the drug treatment, it will have less time to grow, and will likely have a smaller size when the drug reaches the viewing device. We utilized partial least-squares regression (PLSR) to analyze the relations between these functional parameters, as well as their contribution to the ultimate tolerance outcomes.

PLSR is an analytical method that bears close resemblance to principal component analysis. In short, PLSR projects the predictor variables (in our case the quantitative and qualitative characteristics of *M. smegmatis*) and a response variable (in our case, the qualitative drug tolerance outcome) to a new space, creating a linear combination of loading coefficients that approximate the original variables.

The following predictor (X) variables were analyzed:

- Accelerator / Alternator status (qualitative, 1 for Acc, 0 for Alt)
- Length at birth
- Length at drug treatment start
- Birth time relative to drug treatment start, i.e. cell age
- Instantaneous cell growth (from the previous recorded frame)
- Average growth rate since birth

- Cell cycle at drug treatment start (qualitative, 2 for B, 3 for C, 4 for D, 5 for E)
- Start of C cycle relative to drug treatment start
- Concentration of the drug

There was only one qualitative response (Y) variable, treatment outcome (0 - live, 1 - death, 2 - zombie).

Around 28% of variation within the observed variable (training set Y) was explained by the two-component model, and around 25% of variation was predicted by the model according to cross-validation. Adding more components to the model did not improve the fit. The poor fit is most likely a result of noise in the data, and the presence of other unaccounted factors that affected the drug tolerance. However, if one considers the relative simplicity of measured functional parameters (cell length, division timing, and GFP foci presence), even this result can be considered a good first step to deterministic predictions of the drug tolerance outcome.

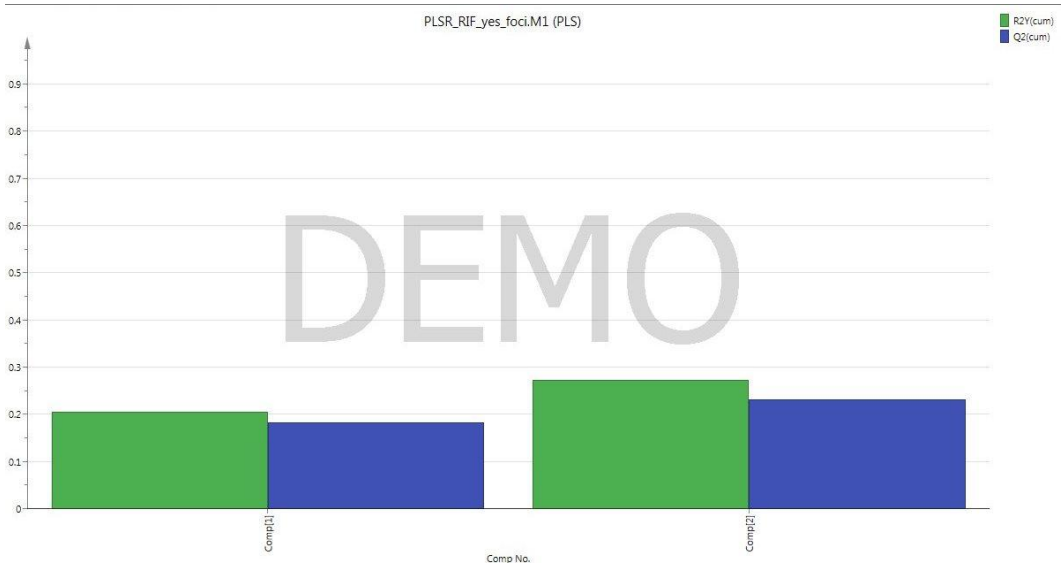


Figure 35: Goodness of Model Fit

Bearing in mind that our model explains slightly more than a quarter of observed drug tolerance response, let us examine it more closely. The score scatter plot showed good clustering of observations with no extreme outliers, and with just three observations outside the tolerance ellipse.

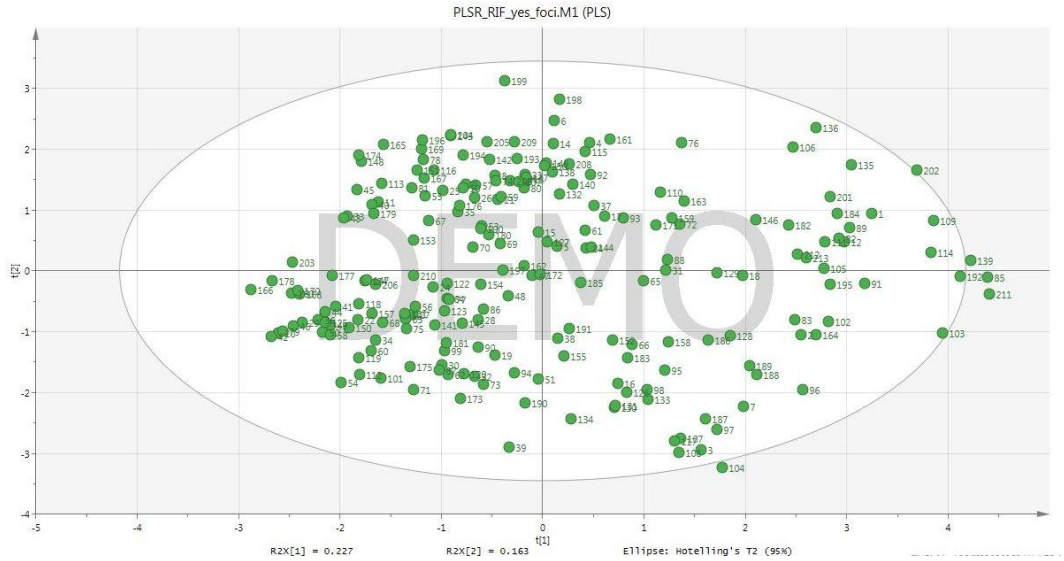


Figure 36: Scores Scatter Plot

The score plot is a map of the observations. It is a window into predictor variable space, which displays how the observations are situated relative to each other.

The model permutation plot indicates that the original model is not spurious, i.e. it doesn't just fits the training set without predicting Y for new observations.

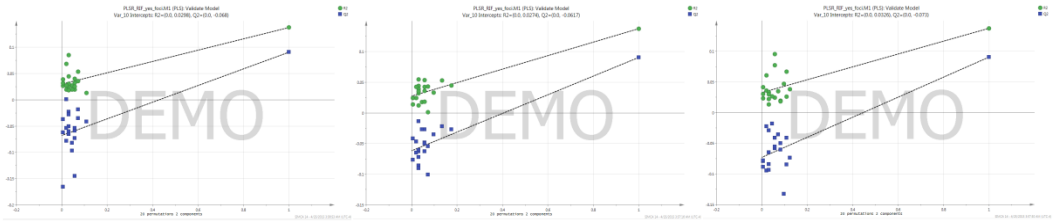


Figure 37: Model Permutation Plots

Three Y-variables were evaluated (from left to right): live, dead, zombie. The values of components for the original model are shown on the right part of each plot, and 20 the Y-permuted models are on the left. According to the SIMCA manual, these graph indicate a valid model: all blue Q2-values to the left are below the original component on the right, and the regression line of the blue Q2 permuted values intersects the vertical axis below zero.

The next plot examines similarity in behavior between variables.

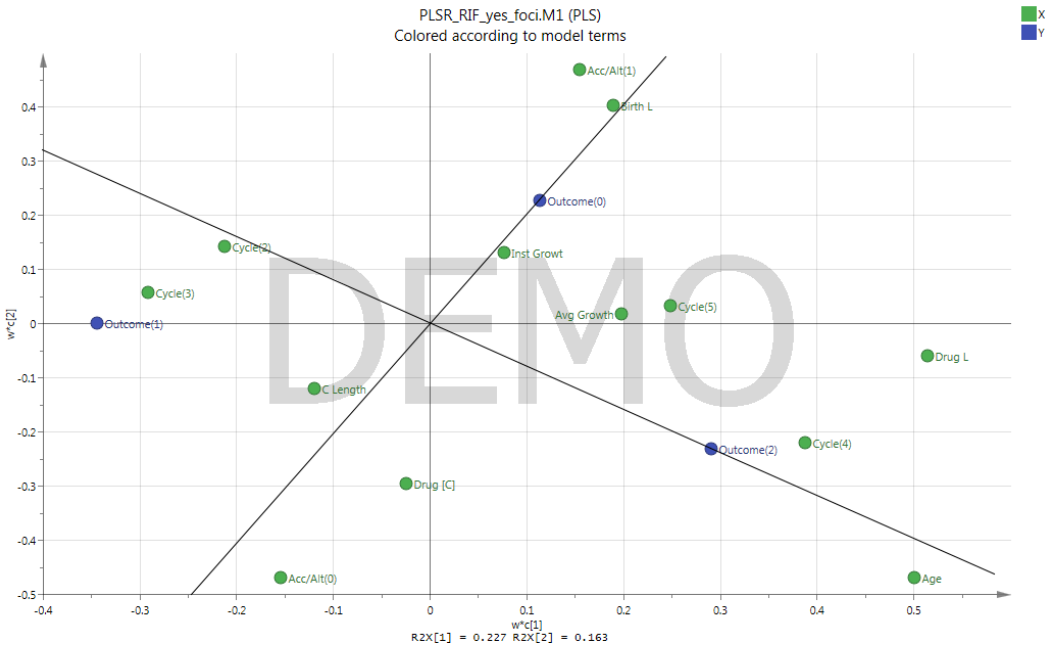


Figure 38: Loadings Weights Plot

To interpret this plot, look at the line that goes through the origin from a Y-variable, and project other X- and Y- variables onto this line at a right angle. Variables that are on the opposite sides of the line are

negatively correlated; values on the same side of the line are positively correlated. The distance between variables represents the strength of correlation.

The next plot presents the correlations between various drug treatment outcomes and the predictor variables from the loadings plot in a clearer format.

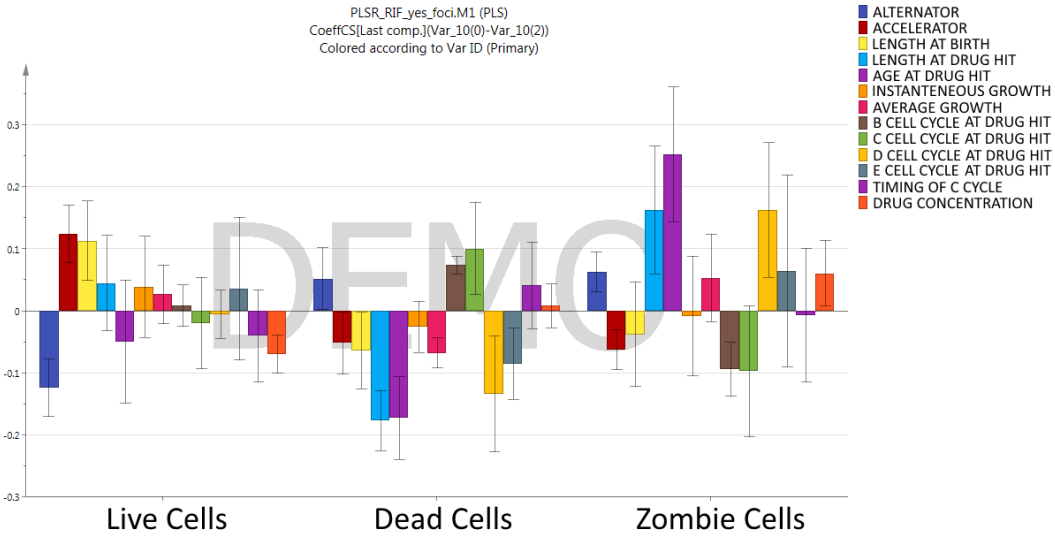


Figure 39: Coefficients Overview Plot

The coefficients are normalized (divided by the standard deviation of their respective observed variable).

As expected, the length and age of the cell at drug hit and certain cell cycle stages correlated heavily with the zombie and dead outcome. Based on the loadings plot, for zombies the age at drug hit and the size at drug hit were weakly negatively correlated, while for the dead cells they were weakly positively correlated. Curiously, the live cell outcome didn't correlate as much with these three functional variables, and was most dependent on the accelerator/alternator status, and the length at birth.

Finally, the variable importance for the projection (VIP) plot summarizes the importance of variables both to explain X and to correlate Y.

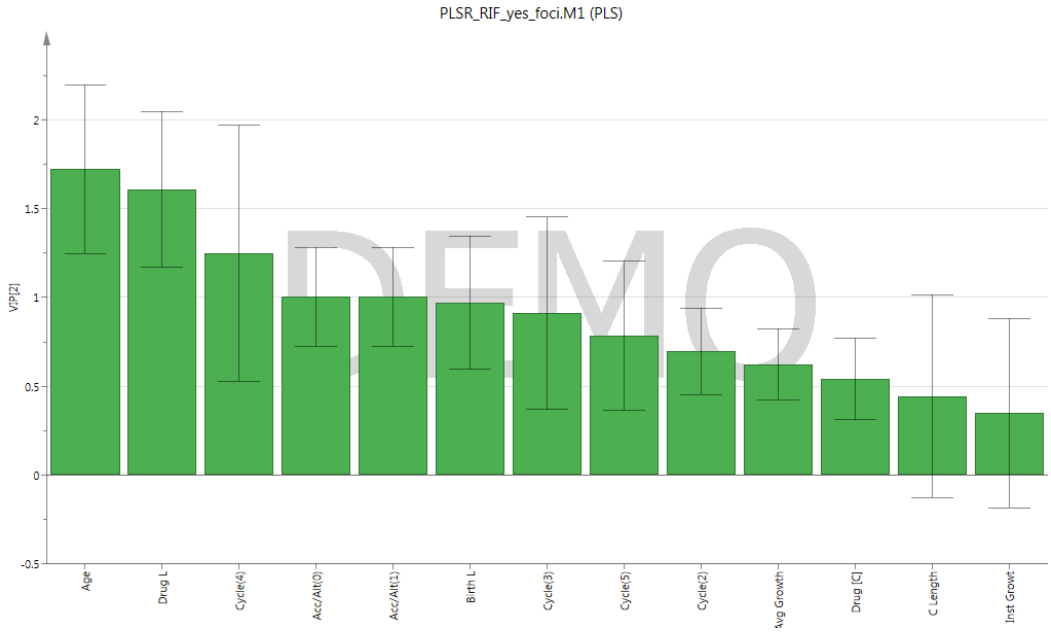


Figure 40: Variable Importance for the Projection

The error bars indicate 95% level confidence interval. Values larger than 1 are considered important, values smaller than 0.5 are considered unimportant, values between 0.5 and 1 are in the "grey zone"

Based on the VIP values, the birth time relative to drug treatment start and cell length at drug treatment start are most important to the rifampicin treatment outcome. Length at birth, D and C cell cycle states at time of drug treatment start, and the accelerator/alternator status also seem important, however the lower boundary of the confidence interval for these values is below the value of one. The instantaneous growth rate right before the drug treatment start, on the other hand, was the least important value in the PLSR model.

Chapter 8. PD Model

8.1 Model Framework

The various functional sub-populations of mycobacteria may be described by using a compartmentalized model. Each compartment represents one distinct sub-populations; all cells within that compartment are assumed to be homogeneous. The rate of change of the number of cells within the compartment is assumed to be either constant, or proportional to the remaining relative cell number in that compartment; in other words it follows zero-order or first-order kinetics, which makes it possible to describe the system with ordinary differential equations (ODEs). The following figure provides the compartmentalized framework example for modeling the drug tolerance within the mycobacteria population.

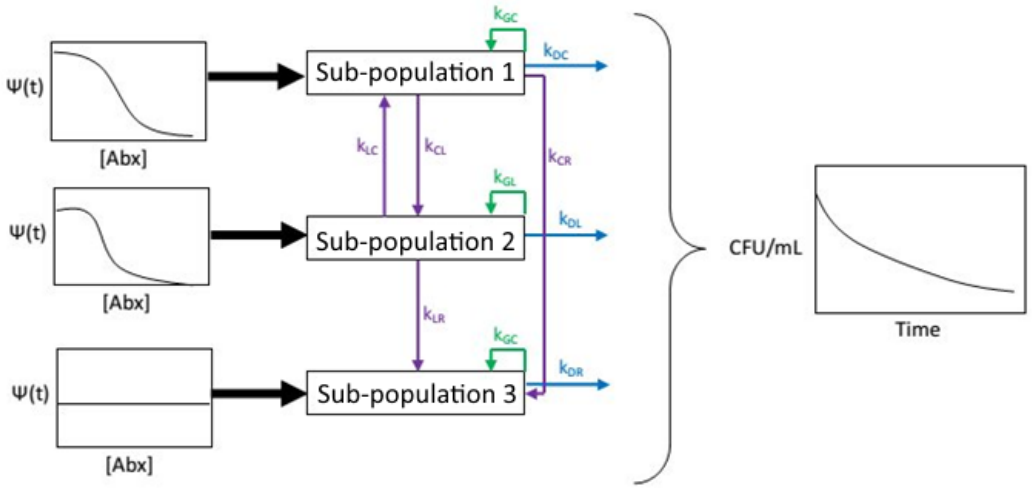


Figure 41: Compartmentalized PD Model

Boxes represent the number of bacteria for each defined phenotype. Arrows represent the first-order transfer rate into and out of particular phenotypic states (such as accelerator/alternator). The model's predictive power should be validated against bulk kill-curve measurements (right).

The PD model[8] consists of the following elements:

1. A set of n bacterial sub-populations denoted by $\sigma = \{1, 2, \ldots, n\}$.

2. State vectors $\mathbf{S}, \mathbf{R} \in \mathrm{R}n$. For $i \in \sigma$, Si represents the number of drug-susceptible, replicating bacteria in sub-population i as a function of time. Ri represents the number of drug-resistant, non-replicating bacteria in sub-population i as a function of time.

3. Rate constants $k, c, kds, kdr \in \mathrm{R}+$, with units of hours^{-1}. kc represents the rate of cell division of replicating cells in the population. kds and kdr represent the rates of cell death among cells in **S** and **R** in the absence of drug.

4. A transition matrix $\mathbf{A} \in \mathrm{R}n \times n$. The ijth entry aij represents the number of cells in sub-population i produced per division of a cell in sub-population j. $\mathbf{a}i$ denotes the ith row of A. Any aij need not be an integer but each column of **A** should sum to 2, since two new cells are produced every division. The rate of change of each Si due to cell growth alone (no S-R switching or cell death) is given by $kc(\mathbf{a}i - \mathbf{e}i) * \mathbf{S}$, where $\mathbf{e}i$ is the ith standard basis vector in Rn. The $\mathbf{e}i$ term represents the "loss" of the parent cell in Si each cell division.

5. A vector $\mathbf{DRUG} \in \mathrm{R}n$. For $i \in \sigma$, $DRUGi$ represents the rate (hours^{-1}) at which a particular concentration of antibiotic kills cells in Si according to an $Emax$ model, which was discussed in Section 2.3

6. A constant $Bmax$ that represents the upper bound for the overall cell population size $B = \mathrm{P}i \in \sigma\ Si + Ri$, and functions Zsr and Zrs that map $\mathrm{R}+ \rightarrow \mathrm{R}+$. Without loss of generality, the rate at which cells in Si become cells in Ri is given by $ZsrSi$, and the rate at which

[8] My initial PD model was based on the (Nielsen, Cars, & Friberg, 2011) paper. I would like to thank Robert Crutcher for improving my first model, and for formalizing its description in a mathematically rigorous way.

cells in *Si* become cells in *Ri* is given by *ZrsRi*. Both *Zsr* and *Zrs* are functions of *B* and *Bmax*. The functions used in this model are

$$Z_{sr} = \alpha_{sr} \ln\left(1 - \frac{B}{B_{max}}\right)$$

and

$$Z_{rs} = \alpha_{rs} \ln\left(\frac{B}{B_{max}}\right)$$

where $\alpha sr, \alpha rs < 0$. These functions are chosen for their property that as $B \rightarrow Bmax$, $Zsr \rightarrow \infty$ and $Zrs \rightarrow 0$, and as $B \rightarrow 0$, $Zsr \rightarrow 0$ and $Zrs \rightarrow \infty$.

For any $i \in \sigma$, the pharmacodynamic differential equations for *Si* and *Ri* are given by:

$$\frac{dS_i}{dt} = k_c(\mathbf{a}_i - \mathbf{e}_i) * \mathbf{S} - (k_{ds} + DRUG_i + Z_{sr})S_i + Z_{rs}R_i$$

$$\frac{dR_i}{dt} = Z_{sr}S_i - (k_{dr} + Z_{rs})R_i$$

A brief word about vector notation: bolded font (**a**) denotes a vector, and *ai* denotes its *i*th element. For **a**, **b** $\in \mathrm{R}n$, **a** $*$ **b** denotes the dot product P*n* *i*=1 *ai* $*$ *bi*. For $i \in \{1, 2, \ldots,$ *n*}, **e***i* $\in \mathrm{R}n$ denotes the *i*th standard basis vector i.e. *e*

ij = 1 if *i* = *j* and 0 otherwise.

8.2 PD Test Case

Let σ = {1,2,3,4}, where each sub-population has the following characteristics:

- i = 1: short, slow-growing cells
- i = 2: long, slow-growing cells
- i = 3: short, fast-growing cells
- i = 4: long, fast-growing cells

Suppose that the daughters of all slow-growing cells are short cells, and the daughters of all fast-growing cells are long cells. Also, every cell gives rise to one slow-growing cell and one fast-growing cell. The transition matrix A for this scenario is:

1	1	0	0
0	0	1	1
1	1	0	0
0	0	1	1

Let us set the rate of cell division (Kc) to 1.46 hours^{-1}, and rates of cell death for both populations (Kds, Kdr) to 0.187 hours^{-1}. For drug action, let us use the following parameters:

i	1	2	3	4
Emax	2.7	3.82	4.38	5.5
EC50	0.00531	.0033	0.0045	0.0025
γ	1.06	1.06	1.06	1.06

Table 5: Training Set Drug Action Parameters

The custom MATLAB code written to implement this PD model produces the following time-kill curve:

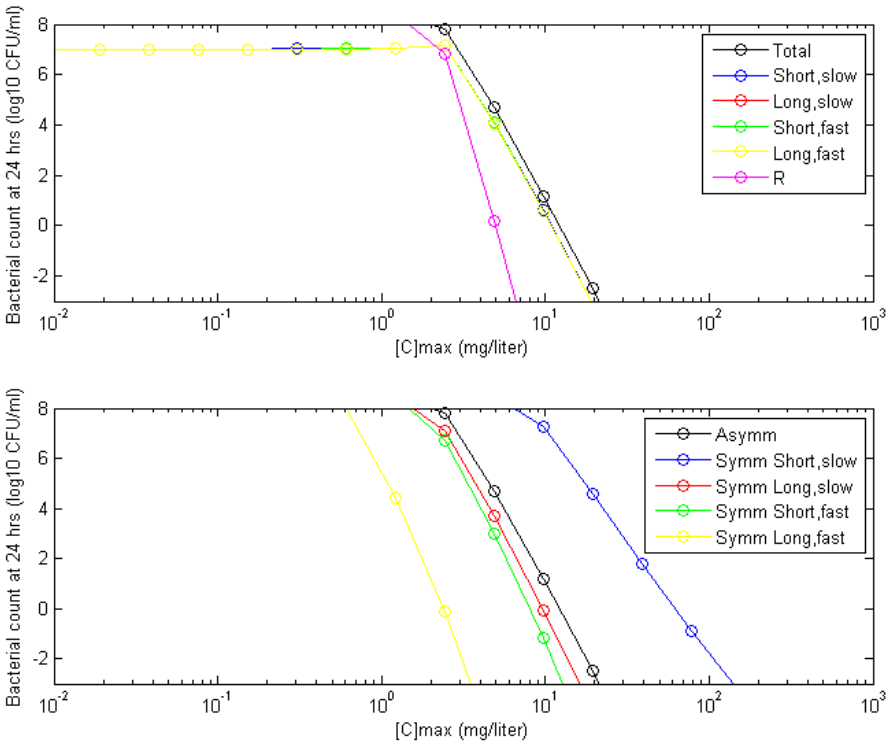


Figure 42: PD Model Training Set, Drug Treatment Outcome

The top plot depicts the MIC curve for bacteria exhibiting asymmetric growth, with each Si compartment depicted, along with the sum of the Ri compartments. The bottom plot compares the total asymmetric cell population to homogenous cell populations with an equal number of initial cells and the same Emax parameters as an individual sub-population .

Chapter 9. Bulk Measurements

In order to assess the usefulness of the PD model framework discussed in the previous chapter, there must be a way to collect bulk *M. smegmatis* population response data. Normally the time-kill curves are collected by exposing bacteria in the exponential growth phase to a certain drug concentrations and monitoring the changes in CFU counts of viable cells over time.

However, in our case the bulk bacterial population must undergo the same treatment protocol as in the live-cell experiments: be grown for ten hours, be subjected to a certain drug concentration for six hours, and then be allowed to recover in drug-free 7H9 media for ten more hours.

We could not find an appropriate commercial-off-the-shelf solution, which would allow us to run this drug protocol in a bulk bacterial population. Therefore, we designed and produced a custom PDMS constant flow device for population measurements.

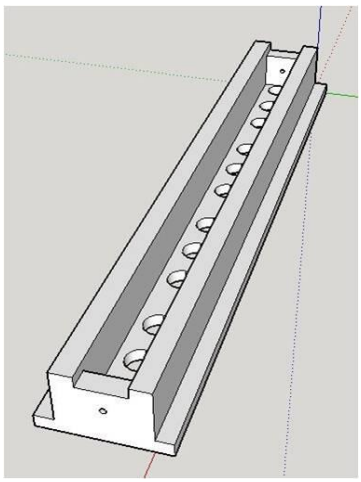


Figure 43: Constant Flow Device for Population Measurements

This device is placed on top of a regular 96-well plate. A strip of filter paper is placed between the device and the wells, to prevent the bacteria from floating out of the wells,

thus lessening the CFU count values. Two media syringe pump feel into one end of the constant flow device via a serpentine mixing device, which allows the user to vary the composition of the media flowing along the main channel of the device, adding and removing the drug. The waste is eliminated out of the other end of the channel, through a plastic tube with 6mm inner diameter. Cells can be kept in a suspended state by placing the 96-well plate with the device on top of a planetary shaker.

Bulk CFU measurements can be taken by carefully piercing the filter paper above a specific well, collecting the cell sample out of that well, and then diluting and culturing it on a 7H10 plate. The device allows up to 12 separate CFU measurements to be taken at various times, before all the wells are exhausted.

A prototype device was created and tested with a .45μm Millipore cellulose filter. Despite some yet unresolved problems with media leakage, the prototype confirmed the feasibility of using a constant flow device for CFU bulk measurements.

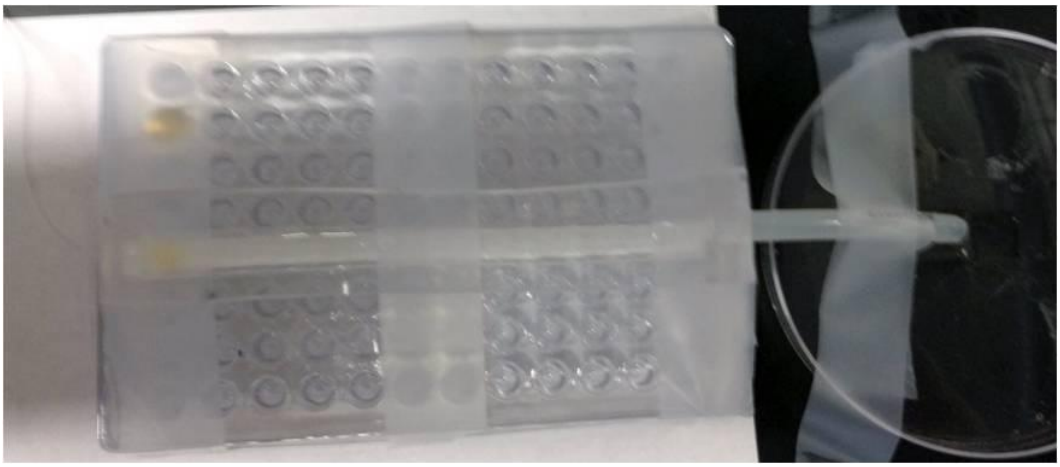


Figure 44: Prototype Constant Flow Device, Top View

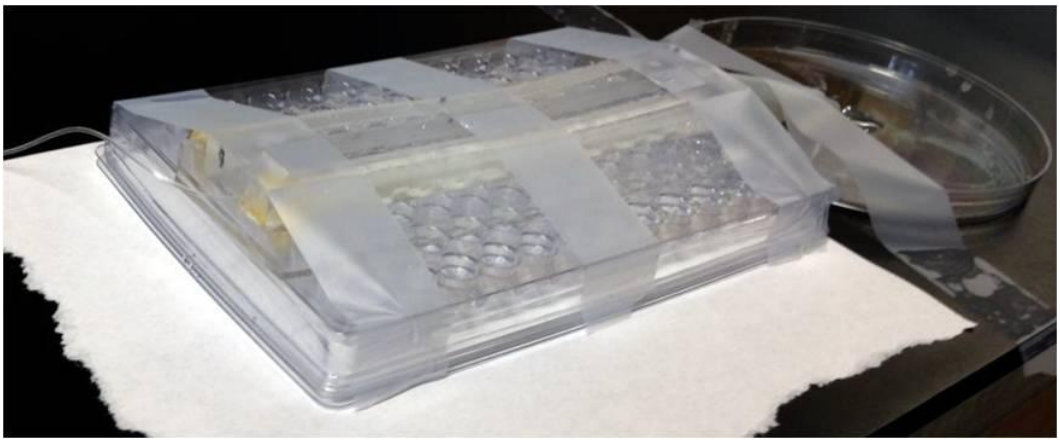


Figure 45: Prototype Constant Flow Device, Side View

Chapter 10. Summary and Discussion

In this thesis we investigated the effects of functional variability of *M. smegmatis* on rifampicin tolerance. It has previously been demonstrated that mycobacteria colonies exhibit distinct heterogeneous phenotypes (namely, accelerator and alternator), and that this functional variation creates sub-populations of cells with differential susceptibility to antibiotics.

With the use of live-cell microscopy system that combines imaging, microfluidics, and computational image processing we separated a *M. smegmatis* population into sub-populations with distinct phenotypic characteristics, and evaluated their response to rifampicin. We first confirmed that there was differential tolerance of accelerators and alternators to rifampicin. We then analyzed the following functional characteristics of mycobacteria: length, age and cell cycle stage of the cell at the start of drug treatment, cell birth size, and growth rate.

Our analysis indicates that a genetically homogeneous M. smegmatis populations can be split into sub-populations with significantly different functional characteristics, and varying antibiotic susceptibility. The cells that died tended to be small at birth, and were generally small, young, and in early stages of their cell cycle when the drug treatment began. The cells that had non-viable offspring (i.e. zombie) tended to be medium-sized at birth, and were generally large, old, and in the late stages of their cell cycle during the beginning of drug treatment. The drug-tolerant cells tended to be large at birth, and at the beginning of drug treatment their population's age and size tended to be in-between that of dead and zombie cells. Incidentally, these results suggest that combining dead and zombie cell population into a single catch-all category would produce a susceptible

population which would have more similar characteristics to the tolerant population that either the dead or the zombie sub-populations, and would thus produce less significant results.

We then conducted PLSR analysis to determine the relative importance of various functional parameters to the rifampicin tolerance. The model fit was rather poor, only around 28% of variation within the observed variable (training set Y) was explained by the two-component model, and around 25% of variation was predicted by the model according to cross-validation. Therefore, out data suggests there are other significant parameters that are not being measured yet. However, given the relative simplicity of measured functional parameters (cell length, division timing, and GFP foci presence), even capturing 28% of variation can be considered a good first step towards achieving deterministic predictions of the drug tolerance outcome based on functional mycobacteria characteristics. As expected, the length and age of the cell at drug treatment start and certain cell cycle stages correlated heavily with the zombie and dead outcome. However, the live cell outcome didn't correlate as much with these three functional variables, and was most dependent on the accelerator/alternator status, and the length of the cell at birth. When the importance of predictor values was evaluated for all three possible treatment outcomes, the birth time relative to drug treatment start and cell length at drug treatment were most important to the rifampicin treatment outcome. The instantaneous growth rate right before the drug treatment start, on the other hand, was the least important value in the PLSR model.

Finally, we created a compartmentalized model framework that is capable of describing the pharmacodynamics of distinct bacterial sub-populations, and which could be used to quantitatively resolve single-cell and population-level measurements.

Chapter 11. Future Directions

11. 1 Further Quantifying the Contribution of Chosen Phenotypic Parameters to Antibiotic Tolerance

Additional data should be collected for better quantification of the contribution of phenotypic parameters to rifampicin tolerance. Ideally, more than 20 microcolonies should be analyzed for each drug concentration, and more rifampicin concentrations should be added (e.g. 4xMIC).

Further sub-population delineation might prove useful, for instance live cells can be split into a sub-population with two live daughter cells, and one live and one dead daughter cell.

Since microcolonies are genetically homogeneous, cross-reference analysis of the difference between microcolonies might help to tease out more details on the contribution of functional parameters to drug resistance.

Drugs other than rifampicin should be investigated, in particular isoniazid.

11.2 Populating the PD Model

The improved single-cell data should be used to populate the PD model (defined in Chapter 8). The model parameters can be estimated using linear regression fitting methods. The output of the model should be compared to the bulk CFU measurements, thus further investigating the influence of the single cell-driven functional mycobacteria variability on antibiotic tolerance.

11.3 Creating a Better Determinant of Drug-Induced Death

As mentioned in Section 5.7, the current method of assigning cells to the "dead" category is far from perfect: absence of growth after the ten hour recovery time does not per se indicate cell death.

Fluorocine diacetate (FDA) could be used as a mycobacteria vitality reporter. At the end of the drug recovery time period, nonpolar (and thus non-fluorescing) FDA could be introduced into the viewing device. Once it is taken up by mycobacteria, it will be hydrolyzed by an enzyme acetylesterase. Live cells will accumulate the polar fluorescein, thus becoming fluorescent (Lawn & Nicol, 2014). However, with FDA vitality reporting the SSB-GFP reporter cannot be used. Additionally, prior to running live-cell FDA images a bulk OD study should be conducted.

Inducible gene expression based upon Tet repressor might provide another possible way for creating a live-death reporting *M. smegmatis* model (Bertram & Hillen, 2008). As with FDA, tetracyclin could be introduced into the viewing device towards the end of the drug recovery period. After tetracyclin is taken up by the mycobacteria, it will inducibly express GFP using a TetON (tetracycline-inducible) promoter, therefore making live cells fluoresce (Martin, 2013).

Chapter 12. Bibliography

Aldridge, B. B., Fernandez-Suarez, M., Heller, D., Ambravaneswaran, V., Irimia, D., Toner, M., & Fortune, S. M. (2012). Asymmetry and Aging of Mycobacterial Cells Lead to Variable Growth and Antibiotic Susceptibility. *Science*, *335*(6064), 100–104. http://doi.org/10.1126/science.1216166

Bertram, R., & Hillen, W. (2008). The application of Tet repressor in prokaryotic gene regulation and expression. *Microbial Biotechnology*, *1*(1), 2–16. http://doi.org/10.1111/j.1751-7915.2007.00001.x

Connolly, L. E., Edelstein, P. H., & Ramakrishnan, L. (2007). Why is long-term therapy required to cure tuberculosis? *PLoS Medicine*, *4*(3), 435–442. http://doi.org/10.1371/journal.pmed.0040120

David, H. L. (1970). Probability distribution of drug-resistant mutants in unselected populations of Mycobacterium tuberculosis. *Applied Microbiology*, *20*(5), 810–814.

Franzblau, S. G., Witzig, R. S., Mclaughlin, J. C., Torres, P., Madico, G., Hernandez, A., … Gilman, R. H. (1998). Rapid, low-technology MIC determination with clinical Mycobacterium tuberculosis isolates by using the microplate Alamar Blue assay. *Journal of Clinical Microbiology*, *36*(2), 362–366.

Hett, E. C., & Rubin, E. J. (2008). Bacterial growth and cell division: a mycobacterial perspective. *Microbiology and Molecular Biology Reviews : MMBR*, *72*(1), 126–156, table of contents. http://doi.org/10.1128/MMBR.00028-07

Jackson, M., Mcneil, M. R., & Brennan, P. J. (2014). NIH Public Access, *8*(7), 1–32. http://doi.org/10.2217/fmb.13.52.Progress

Lawn, S. D., & Nicol, M. P. (2014). Dead or Alive: Can Viability Staining Predict Response to Tuberculosis Treatment? *Clinical Infectious Diseases*, *60*, 1196–1198. http://doi.org/10.1093/cid/ciu1156

Martin, C. J. (2013). Efferocytosis is an Innate Antibacterial Mechanism of Mycobacterium tuberculosis. Retrieved from http://nrs.harvard.edu/urn-3:HUL.InstRepos:10054147

Müller, M., Dela Peña, A., & Derendorf, H. (2004). Issues in Pharmacokinetics and Pharmacodynamics of Anti-Infective Agents: Distribution in Tissue. *Antimicrobial Agents and Chemotherapy*, *48*(5), 1441–1453. http://doi.org/10.1128/AAC.48.5.1441-1453.2004

Nielsen, E. I., Cars, O., & Friberg, L. E. (2011). Pharmacokinetic/Pharmacodynamic (PK/PD) indices of antibiotics predicted by a semimechanistic PKPD model: A step

toward model-based dose optimization. *Antimicrobial Agents and Chemotherapy*, *55*(10), 4619–4630. http://doi.org/10.1128/AAC.00182-11

Nielsen, E. I., Viberg, A., Löwdin, E., Cars, O., Karlsson, M. O., & Sandström, M. (2007). Semimechanistic pharmacokinetic/pharmacodynamic model for assessment of activity of antibacterial agents from time-kill curve experiments. *Antimicrobial Agents and Chemotherapy*, *51*(1), 128–136. http://doi.org/10.1128/AAC.00604-06

Rasband, W. S. (U. S. N. I. of H. (n.d.). ImageJ. Retrieved from http://imagej.nih.gov/ij/

Regoes, R. R., Wiuff, C., Zappala, R. M., Kim, N., Baquero, F., Levin, B. R., & Garner, K. N. (2004). Pharmacodynamic Functions : a Multiparameter Approach to the Design of Antibiotic Treatment Regimens Pharmacodynamic Functions : a Multiparameter Approach to the Design of Antibiotic Treatment Regimens. *Antimicrobial Agents and Chemotherapy*, *48*(10), 3670–3676. http://doi.org/10.1128/AAC.48.10.3670

Sukumar, N., Tan, S., Aldridge, B. B., & Russell, D. G. (2014). Exploitation of Mycobacterium tuberculosis Reporter Strains to Probe the Impact of Vaccination at Sites of Infection. *PLoS Pathogens*, *10*(9), e1004394. http://doi.org/10.1371/journal.ppat.1004394

Vischer, Norbert, Nastase, S. (University of A. (n.d.). ObjectJ. Retrieved from http://simon.bio.uva.nl/objectj/index.html

WHO. (2014). Global tuberculosis report 2014 (WHO/HTM/TB/2014.08). http://doi.org/WHO/HTM/TB/2014.08

Xia, Y., & Whitesides, G. M. (1998). Soft lithography. *Annual Review of Material Science*, *28*(12), 153–184. http://doi.org/10.1146/annurev.matsci.28.1.153