

PHENOTYPIC ALTERATIONS IN *BORRELIA BURGENDORFERI* AND IMPLICATIONS FOR THE
PERSISTENT CELL HYPOTHESIS

AN ABSTRACT
SUBMITTED ON THE TWENTY-FIFTH DAY OF NOVEMBER 2014
TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF MEDICINE
OF TULANE UNIVERSITY
FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
BY

John Russell Caskey

APPROVED: _____
Monica E. Embers, Ph.D., Committee Chair

Mario Philipp, Ph.D.

Andrew MacLean, Ph.D.

James B. McLachlan, Ph.D.

Lucia C. Freytag, Ph.D.

Deepak Kaushal, Ph.D.

ABSTRACT

Lyme disease is the most commonly reported vector-borne disease in the United States. The causative agent of Lyme disease, can alter gene expression to enable survival in a diverse set of conditions, including the tick midgut and the mammalian host. External environmental changes can trigger gene expression in *B. burgdorferi*, and the data demonstrate that *B. burgdorferi* can similarly alter gene expression as a stress-response when it is treated with the antibiotic doxycycline. After treatment with the minimum bactericidal concentration (MBC) of doxycycline, a subpopulation can alter its phenotype to survive antibiotic treatment, and to host adapt and successfully infect a mammalian host. Furthermore, our data demonstrate that if a population is treated with the MBC of doxycycline, a subpopulation may alter its phenotype to adopt a state of dormancy until the removal of the antibiotic, whereupon the subpopulation can regrow. We demonstrate that the chance of regrowth occurring increases as a population reaches stationary phase, and present a mathematical model for predicting the probability of a persister subpopulation within a larger population, and ascertain the quantity of a persister subpopulation. To determine which genes are expressed as stress-response genes, RNA Sequencing analysis, or RNASeq, was performed on treated, untreated, and treated and regrown *B.*

burgdorferi samples. The results suggest several genes were significantly different in the treated group, compared to the untreated group, and in the untreated and regrown group compared to the untreated group, including a 50S ribosomal stress-response protein, coded from BB_0786. The appendices discuss the theory and methods that were used in RNA Sequencing (RNASeq) analysis, and provide an overview of the database that was created for the *B. burgdorferi* transcriptome. Additional studies may demonstrate further how persister subpopulations form, and which genes can trigger a persister state in *B. burgdorferi*.

PHENOTYPIC ALTERATIONS IN *BORRELIA BURGENDORFERI* AND IMPLICATIONS FOR THE
PERSISTENT CELL HYPOTHESIS

A DISSERTATION
SUBMITTED ON THE TWENTY-FIFTH DAY OF NOVEMBER 2014
TO THE GRADUATE PROGRAM IN BIOMEDICAL SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF MEDICINE
OF TULANE UNIVERSITY
FOR THE DEGREE
OF
DOCTOR OF PHILOSOPHY
BY

John Russell Caskey

APPROVED: _____
Monica E. Embers, Ph.D., Committee Chair

Mario Philipp, Ph.D.

Andrew MacLean, Ph.D.

James B. McLachlan, Ph.D.

Lucia C. Freytag, Ph.D.

Deepak Kaushal, Ph.D.

UMI Number: 3680987

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3680987

Published by ProQuest LLC (2015). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

ACKNOWLEDGEMENTS

This dissertation is the lookout for higher mountains that I would not have been able to ascend without support from others. I would like to graciously thank the members of my family for their support and encouragement, especially my mother Carol Bean, and my brothers Anthony and Mike, who continuously offered advice and support. I would like to thank my Aunt Phyllis and Uncle Robert for their continuous help while I lived in New Orleans. I would also like to thank my grandparents Dr. William Bean and Marjorie Bean, who were an inspiration to me, and who paved the way for me to follow. I would like to thank my loving wife, Willow Haley, for her support and encouragement during these long years.

My dissertation advisor, Dr. Monica Embers, has my deepest gratitude and heartfelt thanks for her advice and counsel, which has both aided me to achieve this pinnacle, and enabled me to obtain greater apogees. I would like to thank Dr. Andrew MacLean, Dr. Mario Philipp, Dr. James McLachlan, Dr. Deepak Kaushal, and Dr. Lucia Freytag, for serving on my Dissertation Committee, and

for their guidance and wisdom. I would also like to thank the members of my lab, Nicole Hasenkampf and Amanda Tardo, for their hard work, advice, and assistance during my many experiments.

Many people have offered advice and assistance to me along the way, for which I am extremely grateful, especially Dr. Amanda McGillivray, Bonnie Phillips, Dr. Alejandra Martinez Lege, Avery MacLean, Mary Jacobs, and Michelle and Mason Scarritt.

I would like to acknowledge and thank the LSU Center for Computation and Technology for their hard work in performing the analysis on the RNASeq data.

Last but not least, I would like to thank Dr. Nicholas Quintyne, my undergraduate advisor, who showed me how cool lasers and fluorescence are.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
LIST OF ABBREVIATIONS.....	x
CHAPTER 1: AN INTRODUCTION TO <i>BORRELIA BURGENDORFERI</i>.....	1
1.1 Lyme disease Overview.....	2
1.2 Lyme disease Sequela.....	4
1.3 <i>Borrelia burgdorferi</i> Overview.....	5
1.4 <i>Borrelia burgdorferi</i> Infection Cycle.....	7
1.5 <i>Borrelia Burgdorferi</i> Genome.....	10
1.6 <i>Borrelia burgdorferi</i> Transcriptomics.....	13
1.7 Stochastic Response Mechanisms.....	15
1.8 <i>Borrelia burgdorferi</i> Animal and <i>in vitro</i> Models, and Detection Methods.....	17
1.9 RNA Sequencing (“RNASeq”).....	19

CHAPTER 2: GENERAL INTRODUCTION.....	23
2.1 Purpose.....	23
2.2 Central Hypothesis.....	25
2.3 Introduction and Specific Aims.....	25
2.3.1 Specific Aim 1.....	27
2.3.1 Specific Aim 2.....	27
CHAPTER 3: EFFECT OF ANTIBIOTIC PRESSURE ON PHENOTYPE.....	29
3.1 Introduction.....	29
3.2 Methods.....	31
3.3 Results.....	37
3.4 Discussion.....	49
CHAPTER 4: EFFECT OF BACTERIOSTATIC ANTIBIOTICS ON POPULATION DYNAMICS.....	51
4.1 Introduction.....	51
4.2 Methods.....	53
4.3 Results.....	57
4.4 Discussion.....	65
APPENDICES	
APPENDIX A: MATHEMATICAL MODELING.....	68
1.1 Mathematical Model to Quantify Persisters.....	68
1.2 Description of Differential Gene Expression Analysis.....	73
APPENDIX B: SEQUENCING THEORY AND METHODS.....	74

2.1 Sequencing Background.....	74
2.2 Next Generation Sequencing and RNASeq.....	75
2.3 Next Generation Sequencing Platforms.....	77
2.4 Ion Torrent.....	80
2.4.1 RNASeq on Ion Torrent.....	83
2.4.2 RNASeq Sample Preparation.....	85
2.5 Alignment Algorithms.....	86
2.5.1 Alignment Theory Overview.....	87
2.5.2 Burrows Wheeler Transform.....	90
2.5.3 Seed Matching.....	91
APPENDIX C: RNASeq Database BTech.....	92
3.1 Overview.....	92
3.2 Drupal Overview.....	93
3.3 Adding Content to Drupal.....	94
3.4 Indexing and Searching.....	97
3.5 Displaying Data.....	99
BIOGRAPHY.....	111

LIST OF TABLES

TABLE 1: Table 1. Listing of <i>B. burgdorferi</i> genome, with brief descriptions of <i>B. burgdorferi</i> genes discussed in this dissertation.....	12
TABLE 2: BacLight Staining of Untreated and Treated <i>B. burgdorferi</i>.....	39
TABLE 3: <i>in vivo</i> viability assay results	
TABLE 3A: Results with SCID mice, 3 experiments.....	41
TABLE 3B: Results with C3H mice, 2 experiments.....	41
TABLE 4: Selection of Genes Found to be Significantly Differently Expressed After Treatment with Doxycycline.....	48
TABLE 5: Probability Assay Results	
TABLE 5A: Probability Assay I.....	58
TABLE 5B: Probability Assay II.....	59
TABLE 5C: Comparison of Tubes with Cultures Grown to Population Density and Tubes with Cultures diluted from Stationary Phase....	60
TABLE 6: Comparison of Tubes with Motile Cultures Between Pulse Dose Assay Isolate and B315A19 Isolate.....	62
APPENDICES	
SUPPLEMENTAL TABLE 1: ST1.....	89
SUPPLEMENTAL TABLE 2: ST2.....	90

LIST OF FIGURES

FIGURE 1: Early Staining of <i>B. burgdorferi</i> performed by Dr. Burgdorfer...	3
FIGURE 2: Schematic Diagram of <i>B. burgdorferi</i>	6
FIGURE 3: Life-cycle of <i>Ixodes</i> species Ticks.....	8
FIGURE 4: Overview of BosR and RpoN signaling Pathways, with effects on OspA and OspC.....	14
FIGURE 5: Cases of Lyme disease in the United States reported to the CDC.....	23
FIGURE 6: Cell Count over 5 day doxycycline treatment.....	33
FIGURE 7: Overview of Experimental Process for Mouse Assay.....	34
FIGURE 8: Cell Motility on Day 5 of Assay.....	38
FIGURE 9: BacLight Image of <i>B. burgdorferi</i>	40
FIGURE 10: Mouse RT-PCR.....	42
FIGURE 11: Forest Plot of Gene Upregulation and Downregulation in RNASeq Data.....	44
FIGURE 12: Venn Diagram Showing genes that are significantly expressed above the 2-fold level.....	45

FIGURE 13 A/BTech: Online, searchable database for *Borrelia burgdorferi* genome and transcriptome.....47

FIGURE 14: Representative figure of growth of spirochetes following pulses of doxycycline.....55

FIGURE 15: Pulse Dose Assay Time to Regrowth.....61

APPENDICES

SUPPLEMENTAL FIGURE S1: Plot of k values for mathematical model.....71

SUPPLEMENTAL FIGURE S2: Persister subpopulation regrowth v. time in hours.....71

SUPPLEMENTAL FIGURE S3: Illustration of Sanger Sequencing.....75

SUPPLEMENTAL FIGURE S4: Ion Torrent Sample Processing and Sequencing Overview.....81

SUPPLEMENTAL FIGURE S5: Schematic Overview of the Sequencing Process.....82

SUPPLEMENTAL FIGURE S6: RNASeq Workflow for Sample Preparation.....85

SUPPLEMENTAL FIGURE S7: Fields in Drupal.....94

SUPPLEMENTAL FIGURE S8: Overview of Import Process.....96

SUPPLEMENTAL FIGURE S9: Taxonomy Terms.....98

SUPPLEMENTAL FIGURE S10: Displaying Data on BTech.....100

LIST OF ABBREVIATIONS

ANOVA.....	Analysis of Variance
<i>B. Burgdorferi</i>	<i>Borrelia burgdorferi</i>
BLAST.....	Basic Local Alignment Search Tool
BSK-II.....	Barbour-Stoenner-Kelley-II Media
CDC.....	Center for disease Control
cDNA.....	complementary Deoxy-Ribonucleic Acid
CDS.....	Coding DNA Sequence
CMOS.....	Complementary Metal-Oxide Sensing
CNS.....	Central Nervous System
CRASP-1.....	Complement Regulator-Acquiring Surface Protein 1
CSF.....	Cerebral Spinal Fluid
<i>E. coli</i>	<i>Escherichia coli</i>
ELISA.....	Enzyme-Linked Immunosorbent Assay
emPCR.....	emulsion Polymerase Chain Reaction
HLA.....	Human Leukocyte Antigen
IDS.....	<i>Ixodes Dammini</i> Spirochete

IFA.....Immuno-Fluorescent Assay
MIC.....Minimum Inhibitory Concentration
MBC.....Minimum Bactericidal Concentration
ncRNA.....non-coding Ribonucleic Acid
NGS.....Next Generation Sequencing [Technology]
Osp.....Outer Surface Protein
PGM.....Personal Genome Machine
PTLDS.....Post Treatment Lyme disease Syndrome
rdRNA.....ribosomal-depleted Ribonucleic Acid
RNASeq.....RNA Sequencing
RPKM.....Reads per Kilobase per Million Mapped Reads
RT-PCR.....Reverse Transcriptase-Polymerase Chain Reaction
rRNA.....ribosomal Ribonucleic Acid
SCID.....Severe Combined ImmunoDeficiency
SGS.....Second Generation Sequencing
SNP.....Single Nucleotide Polymorphism

CHAPTER 1: AN INTRODUCTION TO *BORRELIA BURGDORFERI*

1.1 Lyme disease Overview

Lyme Borreliosis, or Lyme disease, has affected humans for over 5000 years. In an effort to determine lineage and glean information from an ancient DNA library, the mummified remains of a Bronze Age human were sequenced using SOLiD sequencing to sequence the whole genome from a bone sample[1]. In addition to genetic polymorphisms, predispositions, and other interesting genomic and phylogenetic data from the human, the researchers found that he had been infected with *Borrelia burgdorferi* (*B. burgdorferi*), the causative agent of Lyme disease [1]. In the 1970's, outbreaks of arthritis in children, which would later be identified as Lyme disease, had been noticed in several areas of the Northeast, particularly in Lyme, CT [2]. An epidemiological effort by doctors and scientists was undertaken to track the symptoms, and establish possible cause, which ultimately led scientist to the *Ixodes* tick and the *B. burgdorferi* spirochete [2].

Currently, Lyme disease is the most commonly reported vector-borne disease in the United States, and cases have risen steadily since reporting

started in 1992 until the current year, 2014. Annual reported cases are round 30,000, but the number may be up to ten times higher due to under-reporting [3]. In the 1970's, Lyme disease was known to begin with an erythema migrans rash, then progress to arthritis, and finally resulting in cardiac and neurological complications [4]. It was initially believed to be transmitted by the *Ixodes* species tick, due to the simultaneous occurrence of tick bites and symptoms, but identifying or culturing a specific pathogen from the suspected tick or from infected patients could not be accomplished [4, 5]. As possible treatments to help afflicted patients, antibiotic therapy with penicillin, tetracycline, and erythromycin were tested for their efficacy after cases of successful antibiotic therapy for Lyme disease in Europe were reported [4]. Penicillin and tetracycline both showed efficacy in reducing patients' symptoms in those studies.

In 1982, Dr. Willy Burgdorfer, for whom the species of spirochete would later be named, hypothesized [6] a connection between infection from ticks harboring a spirochete and Lyme disease. He initially called *B. burgdorferi* the Ixodes Dammini Spirochete (IDS), and published his findings in the *Science* journal [6, 7]. In his paper, he showed Electron Microscopy (EM), immunofluorescence, and Giesma-staining images of the IDS (*B. burgdorferi*) spirochetes. To create the immunofluorescence images, he fed *Ixodes* ticks on rabbits, and then stained the *Ixodes* tick midgut contents with serum from the rabbits (See Figure 1).

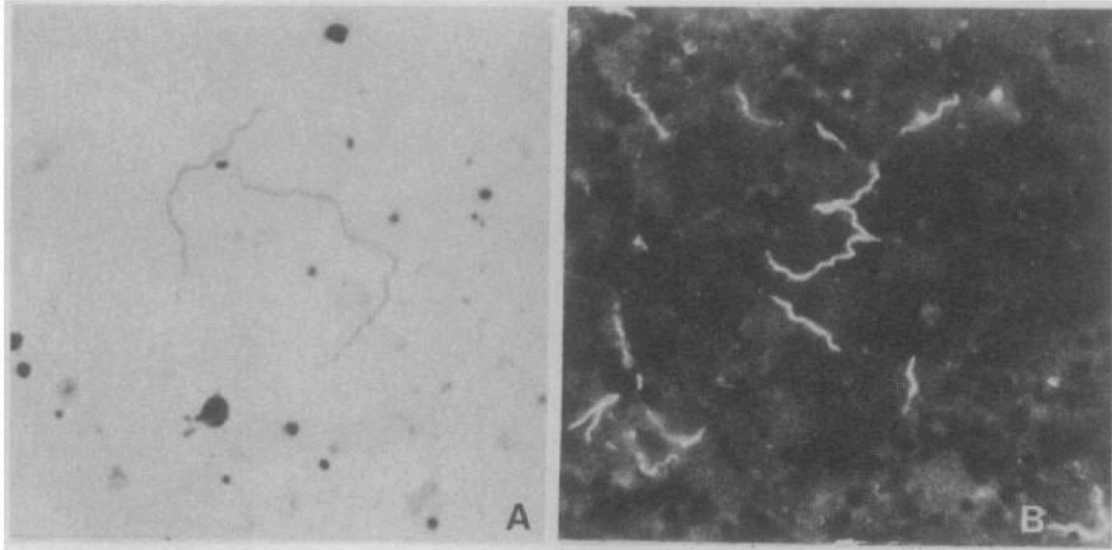


Figure 1. Early staining of *B. burgdorferi* performed by Dr. Burgdorfer. (A) Tick midgut contents are examined and stained with Giesma Staining (B) The serum of an infected patient is viewed using indirect immunofluorescence by tagging spirochetes with antibodies from rabbits previously infected with *Ixodes* species ticks, which had been infected with *B. burgdorferi* [7] Used With Permission.

Dr. Burgdorfer's hypothesis was subsequently confirmed in two ways. The IDS (*B. burgdorferi*) was successfully recovered from the blood, erythema migrans lesions, or the CSF of infected patients and visualized using EM [8]. Patients who suffered from Lyme disease initially experienced elevated IgM antibody titers and subsequently IgG antibody titers, both of which were specific to the IDS (*B. burgdorferi*) [8].

This breakthrough had an enormous effect in the scientific community. Once a causative pathogen had been identified, more experiments could be performed to identify important characteristics of Lyme disease, such as how to treat it, infection progression, and physical characteristics of *B. burgdorferi*. Before these experiments could be accomplished, however, a pressing problem of how to successfully culture it in a laboratory for study was solved in 1984 by Dr. Alan Barbour. Dr. Barbour published a method to create Barbour-Stoenner-

Kelley-II media, which could be used to culture *B. burgdorferi* *in vitro*. A study which was published in 1997, that used Sanger Sequencing to sequence the genome of *B. burgdorferi*, further pushed the capabilities of research forward by allowing research on the entire *B. burgdorferi* genome [9].

1.2 Lyme disease Sequela

Lyme disease is formally divided into early and late Lyme disease. To transmit *B. burgdorferi*, an *Ixodes* species tick must typically take a bloodmeal for at least 36 hours [10, 11]. Early Lyme disease typically presents initially as a distinctive erythema migrans, or “bull’s-eye rash”, although this may present as a nonspecific dermatitis, or not at all [12, 13]. Other symptoms of early Lyme disease include fatigue, malaise, muscle and joint stiffness or pain, and migraines [14-16]. If untreated, *B. burgdorferi* can disseminate to immune-privileged sites, and the central nervous system (CNS), heart, and joints, causing meningitis, carditis, and arthritis, respectively [14, 16]. If Lyme disease progresses to late Lyme disease, neuropathy, cognitive dysfunction, and arthritis, with long-term residual effects may occur [16]. Treatment for Lyme disease usually is doxycycline 200 mg/daily for 14-21 days, or amoxicillin 1500 mg/daily for 14-21 days if doxycycline is contraindicated, and Ceftriaxone 2g/daily for 30 days intravenously is recommended if there is evidence of dissemination to the CNS [16]. A recent study suggests that if *B. burgdorferi* have been able to establish a persistent, niche infection in a patient, third-generation

cephalosporins like ceftriaxone, or the macrolide carbomycin, may be more effective [17].

Lyme disease can easily mimic other commonly acquired diseases such as the flu, which can complicate a prompt and accurate diagnosis and subsequent treatment. To assist in accurate treatment, and to prevent incorrect medicating, the CDC has established criteria for the diagnosis of Lyme disease. Diagnosis of Lyme disease requires a positive ELISA or IFA test, with a positive Western Blot to confirm [18]. Depending on the timing of symptoms, the Western Blot will probe against IgG, or IgM antibodies against *B. burgdorferi* surface proteins. Once a patient receives antibiotic therapy, Lyme disease usually resolves, but in some instances of Lyme disease, patients either experience a relapse of symptoms, or the symptoms never fully resolve, which is formally known as post treatment Lyme disease syndrome (PTLDS). PTLDS is highly controversial, and there is no clear consensus on the appropriate treatment regimen to help patients who are still experiencing symptoms after the conclusion of antibiotic therapy.

1.3 *Borrelia burgdorferi* Overview

B. burgdorferi is an approximately 10-30- μm long, and 0.5- μm wide spirochete in phylum *Spirochaetes*, class *Spirochaetes*, order *Spirochaetes*, family *Spirochaetaceae* [19]. *B. burgdorferi* is exclusively transmitted by the *Ixodes* species tick. In North America, *Ixodes scapularis*, and *Ixodes pacificus* are the primary tick vectors, while in Asia and Europe, *Ixodes ricinus* and *Ixodes*

persulcatus are the tick vectors, while most recently, in South America, *Ixodes pararicinus* was found to be a tick vector [6, 19, 20]. If a *B. burgdorferi* is viewed from the exterior to the interior, it has a surface membrane on the exterior, beneath that is a periplasmic space where the flagella bundle runs, beneath that is a peptidoglycan layer, beneath that is a cytoplasmic membrane, and finally at the interior is the protoplasmic cylinder, where nuclear and metabolic machinery is housed (See Figure 2) [19, 21-23].

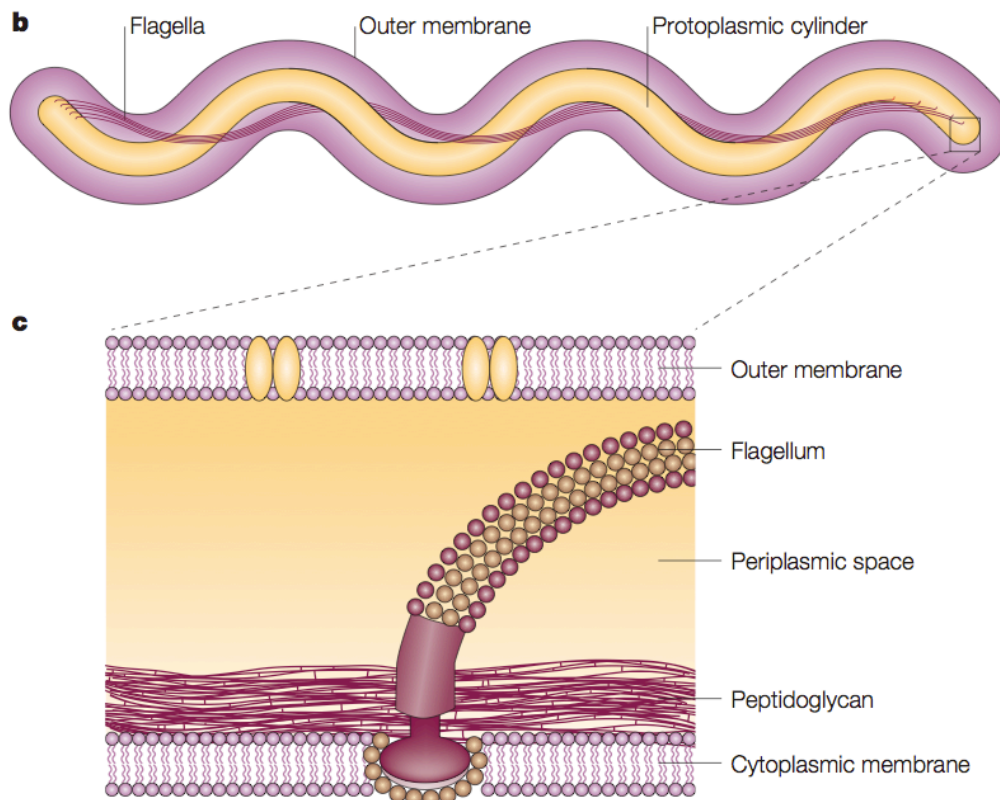


Figure 2. Schematic Diagram of *B. burgdorferi*. On its exterior facing surface, *B. burgdorferi* has a cytoplasmic membrane, an internalized flagellum anchored to the membrane, a peptidoglycan layer, periplasmic space, and an outer membrane [19] Used With Permission.

Although many of the specific functions of the outer surface proteins are still unknown, it has been shown that *B. burgdorferi* can alter expression of some of

its surface proteins, such as outer surface protein A, abbreviated OspA, and outer surface protein C, abbreviated OspC, to survive within the *Ixodes* species tick, and to evade immune detection [24]. Within the periplasmic space, the flagellum is composed of a bundle of 7-11 flagella, arranged in a left-handed helix, which drives rotation counter-clockwise as the spirochete moves [22, 25].

Staining and viewing *B. burgdorferi*, like most spirochetes, cannot be reliably accomplished with a simple Gram-stain, although technically it does stain weakly gram-negative due to its thin, internal peptidoglycan layer. It can be viewed with fluorescence microscopy by tagging surface proteins (usually OspA) with an antibody, electron microscopy, darkfield microscopy, and Giemsa staining [6].

1.4 *Borrelia burgdorferi* Infection Cycle

B. burgdorferi maintains a complex lifecycle for transmission and infection. Since trans-ovarian infection rarely occurs, the primary mechanism of transmission is by the *Ixodes* species tick taking a bloodmeal from an infected host and subsequently feeding on an uninfected host [19]. After acquisition by the tick, *B. burgdorferi* resides in the tick midgut, until transmission to a new host [11]. *Ixodes* species ticks will seek bloodmeals from vertebrate animal hosts. This activity is called “questing,” and takes place in the summer for larva, in the following spring and early summer by nymphs after molting, and in the fall by the

adults after a final molting process (See Figure 3) [19, 26].

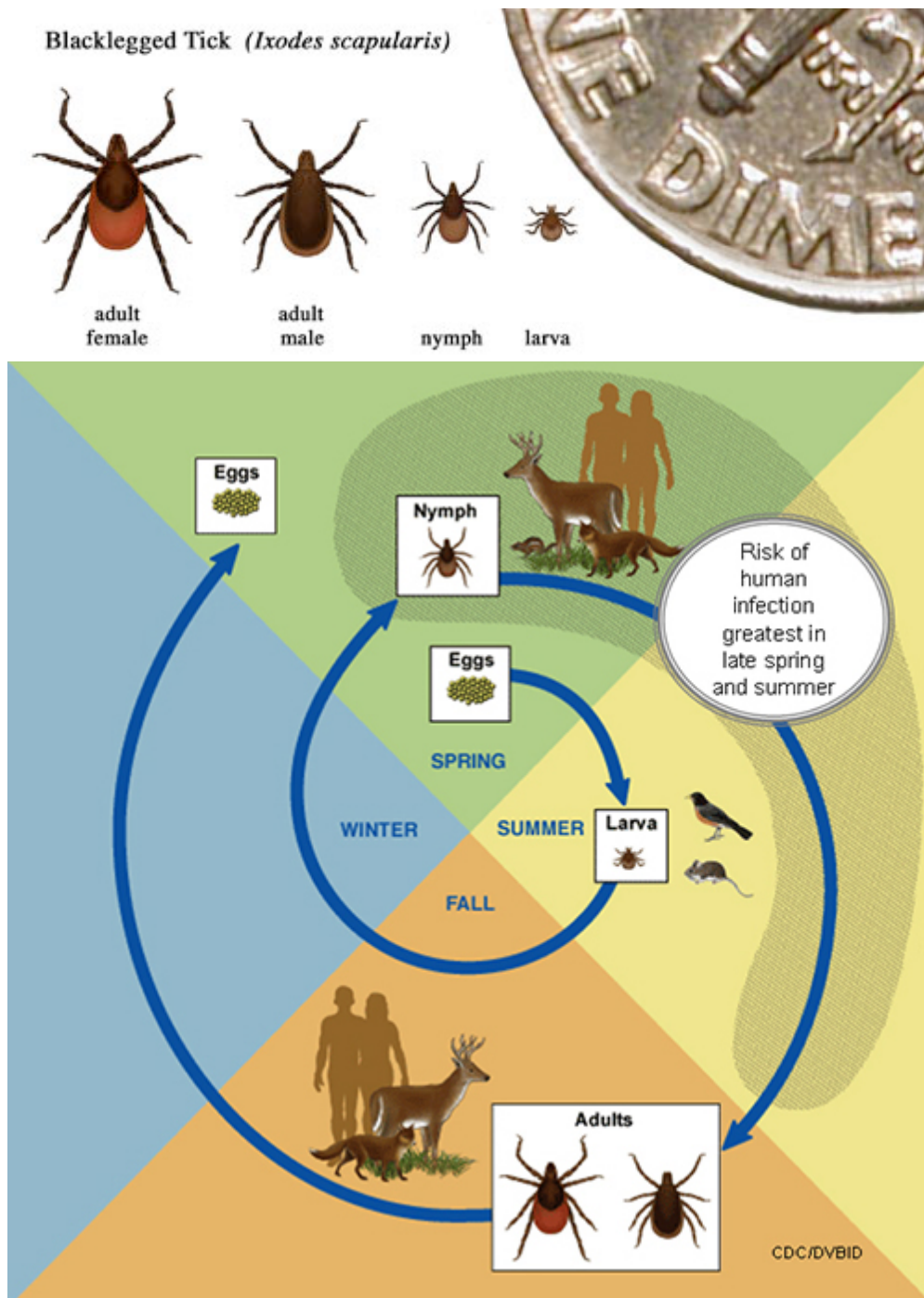


Figure 3. Life-cycle of *Ixodes* species ticks. (A) Relative size of *Ixodes* species ticks. (B) Life cycle and affected animals. Due to their small size and emergence during peak times of activity, nymphs pose the greatest threat [18]. Used with Permission.

Interestingly, small reptiles such as lizards appear to impede the transmission of *B. burgdorferi* after a bloodmeal by *Ixodes* species ticks, possibly by differences in the host complement system [27, 28]. Typically, a larval or nymphal *Ixodes* species tick will feed from a reservoir host such as the white-footed mouse, *Peromyscus leucopus*, then after molting, will transmit *B. burgdorferi* to a second host when it feeds again [29-31]. Adult *Ixodes* species ticks will usually feed on medium to larger mammals, such as opossums, raccoons, the white-tailed deer *Odocoileus virginianus* [26]. Although *B. burgdorferi* can be transmitted either by nymphs or adults, because of their size and activity during spring and early summer, the transmission of *B. burgdorferi* from nymphs holds the greatest risk of acquiring Lyme disease [31].

B. burgdorferi have numerous outer surface proteins (Osp) which can be either expressed on the exterior of the outer surface membrane, or hidden in the periplasmic space at different stages of the infection cycle to promote survival [21, 24, 32]. OspA and OspB are expressed on spirochetes in the *Ixodes* tick species midgut after initial colonization, but repressed after host mammalian infection [33, 34]. After tick midgut colonization, *B. burgdorferi* can survive for months in a low-nutrient state by selectively expressing several genes, such as *bb0365*, *bb0690*, a Dps homolog, and *bptA*, which alter the *B. burgdorferi* phenotype to survive in a low-nutrient state [35-37]. Environmental factors such as pH and temperature can trigger a change back into a rapid- or active-growth phenotype [38].

1.5 *Borrelia burgdorferi* genome

The *B. burgdorferi* genome is composed of a chromosome, nine linear plasmids, and twelve circular plasmids [19, 39]. Specific plasmid maintenance regions maintain proper fidelity and replication of the *B. burgdorferi* plasmids during replication [40]. These plasmid maintenance regions can also be used to selectively insert or delete genes by the use of shuttle vectors [40-42]. Circular plasmids usually maintain a supercoiled state when not in transcription or replication mode, and can form secondary structures [43]. Linear plasmids and the linear chromosome have telomere structures at the ends with hairpin structures and covalently-closed strands [9, 39, 44]. Replication of linear plasmids and the chromosome is initiated at the center, then extends outwards to the ends, where the ResT protein forms the telomere [45, 46]. If a plasmid is truncated or lost during cell passage, *in vitro* growth may not be affected, but *in vivo* growth and infectivity can be altered as a result, and loss or alteration of plasmids must be monitored during an experiment [47-49].

The chromosome encodes important cellular features such as tRNA and rRNA, however the *B. burgdorferi* organism has evolved a complex system of different genomic elements. For example, ResT assists with the formation of the telomere on linear plasmids and the linear chromosome, but is found on circular plasmid 26 (cp26) [45]. A subset of the circular plasmids, cp32, which may have anywhere from 1-7 members, have many redundant sequences, and have been

hypothesized to originate from a bacteriophage [50-53].

The *B. burgdorferi* genome is still being explored, with many coding sequences (CDS) still only “predicted” or “hypothetical”. Table 1 shows the *B. burgdorferi* genome, with brief descriptions of relevant genes contained in each plasmid or chromosome.

Table 1. Listing of <i>B. burgdorferi</i> genome, with brief descriptions of <i>B. burgdorferi</i> genes discussed in this dissertation.		
Name	Genomic Type	Notes or Notable Genes/Proteins
lp5	linear plasmid	Outer Membrane Protein, hypothetical proteins
lp17	linear plasmid	protein CdsM, other pseudogenes
lp21	linear plasmid	predicted proteins
lp25	linear plasmid	protein p23, required for infectivity[1,2]
lp28-1	linear plasmid	transmembrane protein, required for infectivity[1,2]
lp28-2	linear plasmid	DNA helicase, predicted proteins
lp28-3	linear plasmid	predicted proteins
lp28-4	linear plasmid	virulence-associated lipoprotein
lp36	linear plasmid	fibronectin-binding protein, predicted proteins
lp38	linear plasmid	opp proteins, ospD
lp54	linear plasmid	outer membrane proteins, ospA, ospB
lp56	linear plasmid	ErpX, BppA
cp26	circular plasmid	ospC, opp proteins
cp32-3	circular plasmid	BppB proteins
cp32-4	circular plasmid	predicted proteins
cp32-6	circular plasmid	predicted proteins
cp32-7	circular plasmid	predicted proteins
cp32-8	circular plasmid	predicted proteins
cp32-9	circular plasmid	Erp-family, Bpp-family proteins, predicted proteins
cp9	circular plasmid	predicted proteins
chromosome	cchromosome	tRNA, rRNA, metabolic proteins

1.6 *B. burgdorferi* Transcriptomics

B. burgdorferi can express multiple genes in response to environmental conditions. As a notable example, *ospA* is expressed in the tick midgut and briefly in an infected host, while *ospC* is expressed in the infected host but not in the tick midgut [34, 54, 55]. Expression of these genes follows conditions that would appear in both the tick midgut and a mammal. Temperature, pH, cell density, and expression levels of other *B. burgdorferi* genes all play important roles in the *B. burgdorferi* transcriptome regulation [34, 38, 54-56]. As a result, protein products can be grouped into OspA-like and OspC-like categories, depending on when they are expressed during the *B. burgdorferi* life cycle, and what causes upregulation or downregulation [54].

Generally, both lower temperature and higher pH trigger upregulation of *ospA*-like genes and downregulation of *ospC*-like genes, while higher temperature, lower pH, and higher cell density trigger upregulation of *ospC*-like genes and downregulation of *ospA*-like genes, although it has been well established that *ospA* specifically is not affected by cell density [54, 56]. Recent work suggests OspA-like proteins are expressed constitutively in the tick but not the host, while OspC-like proteins are induced by changes to a host environment and can cause down-regulation of OspA-like proteins [54, 56, 57].

Another pathway that regulates virulence in *B. burgdorferi* is called the RpoN-RpoS pathway (See Figure 4) [58]. Environmental signals trigger

phosphorylation of histidine kinases and phosphorylation of Rrp2 (Response regulatory protein 2), which binds and activates RpoN, which then activates *rpoS* transcription [59-63]. RpoS regulates numerous transcriptional products, including OspC, DbpA, and the fibronectin-binding protein BBK32 [64]. The *rpoS* gene has been hypothesized to play a role in differential gene expression and virulence by inducing expression of *ospC* and *dbpAB* [65]. The BosR (oxidative response regulator) protein has been shown to regulate both RpoS and OspA after successful infection into a host [57].

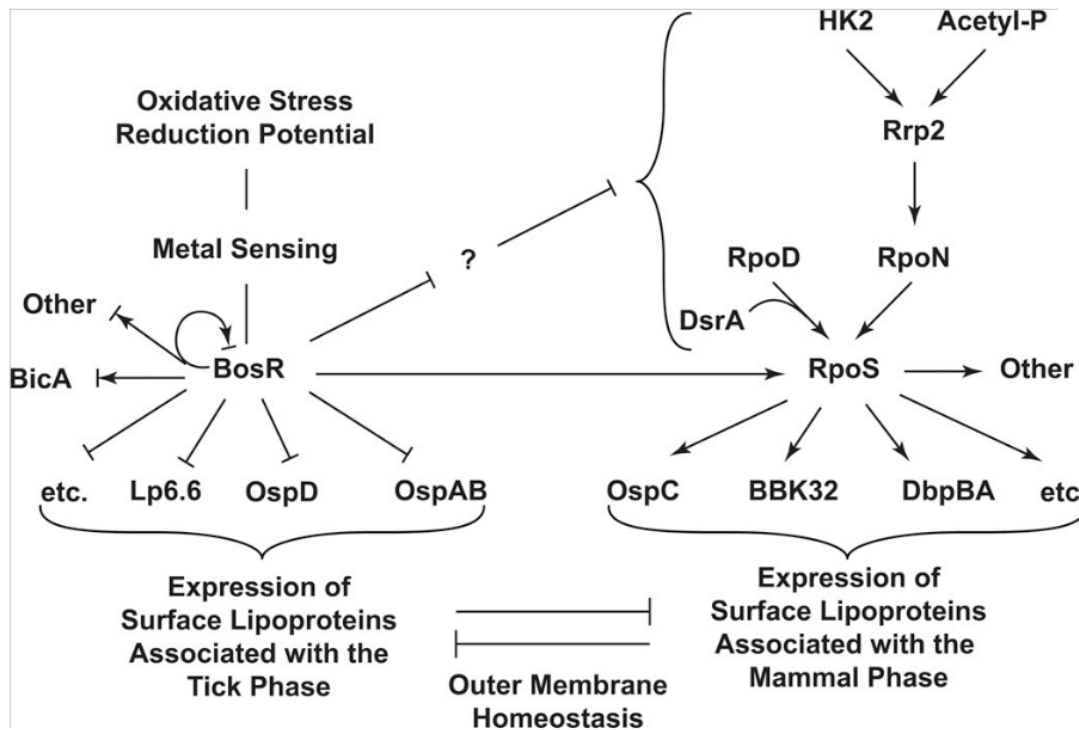


Figure 4. Overview of BosR and RpoN signaling pathways, with effects on OspA and OspC [57]. Used with Permission.

Once inside a mammalian host, *B. burgdorferi* expresses proteins in the OspE family to bind and inactivate host complement factor H [66-68]. This

includes complement regulator-acquiring surface protein 1 (CRASP-1) which is expressed in the host, but repressed in the tick midgut [21, 24, 69]. The capability to inactivate the host alternative complement pathway serves both as a tool for initially suppressing the immune response and breaking a chain in host pathogen detection and elimination and for long term survival and persistence [67]. To evade immune recognition, the lp28-1 plasmid contains multiple *visE* genes, which undergo recombination events to form different external VlsE lipoproteins [68, 70, 71]. The OspC protein is the dominant antigen, and is initially expressed at high levels to promote establishing an infection, but it is subsequently downregulated once the host antibody response develops [70-72].

1.7 Stochastic Response Mechanisms

Since *B. burgdorferi*, like many other bacteria, can change gene expression at different times to respond to stimuli. Determining how and when expression can change is of key interest to researchers in the field. As a more concrete example, it has been shown that a temperature increase can cause OspC expression, along with other host-specific signals, which then results in the down-regulation of OspA [54]. While genetically identical, a bacterial population may not respond to a stimulus in exactly the same way, but instead may gradually respond to changes in temperature, pH, or a host response factor. Intuitively, this makes sense: the temperature will not immediately be 37⁰ C for every bacterium, and the host response factor signaling molecule will not

immediately bind to target receptors in every bacterium. A simplified and commonly used model is to state that the population of bacteria will change expression of OspA to OspC, assuming the bacterial population successfully infects the host. However, this model neglects changes that can take place for individual bacteria and subpopulations, which can affect the overall population dynamics [73, 74]. An individual bacterium within a population should induce OspC and downregulate OspA to survive in the host. If they do not, then it is unlikely that they will survive in the host environment [70], unlike most of the surviving population, which will change expression. Stochastic responses in bacteria can be defined as random fluctuations in a bacterium's gene expression which result in changes from the overall population [75]. In cases of growth and reproduction, stochastic responses can play a less distinct though important role. At a certain point during growth, a bacterium will transition to a reproduction mode, and the two daughter cells will immediately transition to growth. During a given time interval, each bacterium may (or may not) begin division to expand the population, but each replication event has a quantifiable probability of occurring [73, 74].

This modeling can be carried over to other aspects of bacterial metabolism as well. During the *B. burgdorferi* life cycle, it must transition from a tick-survival state to an animal-survival state, and back, to successfully infect an animal host. A successful infection takes approximately 30 hours of feeding from an *Ixodes* species tick, and during a tick bloodmeal, migration does not occur before 6

hours have elapsed [76]. Therefore, there is a random chance that environmental triggers like blood and saliva will induce expression of sufficient response genes within *B. burgdorferi* to activate motility pathways. If this gene expression occurs, the *B. burgdorferi* will begin migration from the midgut to the hypostome of the tick, and finally the host. *B. burgdorferi* may also induce stress-response genes in response to external stresses, and therefore individual *B. burgdorferi* will induce response genes as a result, which may be different from the overall population. The random response of individual bacteria and bacterial subpopulations will create differences in the population, and this can be beneficial if the population must survive a stressful environment.

1.8 *Borrelia burgdorferi* Animal and *in vitro* Models, and Detection Methods

As mentioned previously, the *B. burgdorferi* organism exists in different environments, briefly and for an extended period, throughout its life cycle. It may remain for months in the tick midgut, and must pass through the tick mouth. It must survive within an animal host, until a tick takes a bloodmeal, when the cycle of transmission continues. Reservoir host animals, such as mice and other rodents, usually do not show signs of infection such as arthritis, but can harbor a persistent infection [77]. Other animals like dogs and rabbits may display signs like a rash or arthritis, but maintain a persistent infection or clear the infection, respectively [77]. In a non-human primate model using rhesus macaques, *B. burgdorferi* can successfully maintain a persistent infection, despite antibiotic

treatment [78]. The rhesus macaque also can display an erythema migrans rash, and the signs of central and peripheral nervous system disease [79-81].

Numerous challenges exist for studying *B. burgdorferi* in a laboratory setting. Initially, *B. burgdorferi* was viewed by allowing an *Ixodes dammini* tick infected with *B. burgdorferi* to take a bloodmeal on a rabbit, and then using the antibodies from the rabbit for immunofluorescence [6]. Other animals such as a mouse model may be used to study infection, however, animals besides humans usually do not develop signs characteristic of Lyme disease such as carditis and arthritis, and mice are a natural reservoir host and not an end host. This problem was mitigated by the discovery by Barthold et al that the C3H/He inbred strain displayed arthritis and carditis after infection with *B. burgdorferi* [82]. A non-human primate model can be used to study infection and treatment with *B. burgdorferi* [78] however, high cost is a consistent concern when undertaking studies of this kind. When studying *B. burgdorferi in vitro*, after multiple passages the strain may lose plasmids spontaneously, which can add a confounding factor to the experiment, and must always be monitored during the course of the experiment [40, 47, 83].

In 1984, Dr. Alan Barbour modified an existing media [84, 85] to create one suitable for culturing *B. burgdorferi in vitro*, dubbed Barbour-Stoenner-Kelley-2 media (BSK-II) [86]. BSK-II required rabbit serum to be added, since *B. burgdorferi* are believed to be unable to synthesize some fatty acids found in the

serum [87]. A commercially-available, standardized version, BSK-H, supplemented with rabbit serum is also available.

In a clinical setting, while it is possible to culture *B. burgdorferi* from blood, urine, cerebral-spinal fluid (CSF), and joint synovial fluid from an infected host [7, 88], the CDC recommends employing a two-step diagnostic approach using an ELISA and a western blot instead of attempting to culture specimens [89]. Because of the difficulty culturing *B. burgdorferi* after infection, or growing *B. burgdorferi in vitro* after it has host-adapted, the use of DNA copy number and RNA transcript abundance has been performed to indicate the presence of *B. burgdorferi* [90, 91]. DNA copy number can provide qualitative confirmation if a probe against specific genes is used in a PCR assay [91, 92], while RNA transcript abundance can provide information on the presence of a *B. burgdorferi* infection.

1.9 RNA Sequencing (“RNASeq”)

RT-PCR can be used to determine expression of specific genes, while microarrays and Next Generation Sequencing (NGS) approaches can both determine a sequence and expression on a much larger scale. In NGS and microarray experiments, a normalizing factor is usually used to quantify expression, such as a constitutively expressed gene, or a mathematical normalization factor like Reads per Kilobase per Million Reads (RPKM) [93-95]. The RPKM value normalizes the reads that are mapped to an exon, by all reads

and exons in the experiment. NGS can provide transcript numbers not only for genes of interest, but for groups of genes, and potentially entire transcriptomes of organisms over different experimental conditions, thus enabling the discovery of novel genes, and the quantification of differential expression.

The analysis of which genes in bacteria are expressed in response to a given environmental stimulus [96], or bacterial differential gene expression (differential gene expression in this context), has become an important part of understanding bacterial pathogenesis, virulence, and growth. RNA Sequencing, or “RNASeq” for short, is the application of sequencing methods like NGS to determine a particular sequence of RNA in an organism. RNASeq techniques began with microarrays, but now involve primarily NGS methods on large scale, or high-throughput sequencing, with a variety of purposes. Depending on how the experiment is designed, a few possible applications of RNASeq are identification of noncoding RNA (ncRNA), the amount of a gene that is expressed under different conditions, or identification of novel genes with a previously unknown purpose [93]. To set up the experiment, RNA is extracted from the sample, preprocessed, processed into cDNA, and then sequenced in one of several commercially available sequencers, such as Illumina, Roche, or Life Technologies.

After sequencing is performed, additional challenges remain. The first step is determining how the genomic transcripts from the sample align to the corresponding organism’s reference genome. Ideally, when designing an

experiment, another factor called sequencing depth should also be taken into account. Sequencing depth can be described as how much of the genome needs to be aligned with either RNA (cDNA) or DNA to accurately support or nullify the hypothesis. For example, if every nucleotide in a genome were covered, physical coverage of the genome would be complete, however, if only one read spanned each segment, there would be little indication of how accurate misaligned reads were, whether they were sequencing errors or Single Nucleotide Polymorphisms (SNP) [97]. If multiple reads spanned each segment of the genome, then more information could be obtained from the analysis.

Complicating the analysis is the fact that due to polymerase replication errors or sequencing errors, a given genomic transcript may not exactly match the reference genome, but may instead have any number of mismatches, and a relatively small genome can be over a hundred thousand kilobases long. This task cannot be accomplished by simply attempting to match the genomic transcript against every possible location where it could align, which is formally called a “brute force” method. Instead, tasks of discrete, sequential steps, called algorithms, have been developed to shorten the process, and are used in computer programs to determine the best match a given transcript has in the genome. Once transcript levels have been mapped to a reference genome, gene expression can be quantified, and novel genes can be discovered [93, 98]. If one or more conditions are tested in an experiment, NGS can provide a highly accurate way to determine differential gene expression [99-101]. For a full

description of both RNASeq and Sequencing Theory, and a brief overview of the theory behind creating alignment algorithms, see Appendix B.

CHAPTER 2: GENERAL INTRODUCTION

2.1 Purpose

Borrelia burgdorferi, the causative agent of Lyme disease, is the most common vector-borne pathogen in the United States [102]. The number of cases, reported and unreported, has increased steadily over the past decade (See Figure 5) [102].

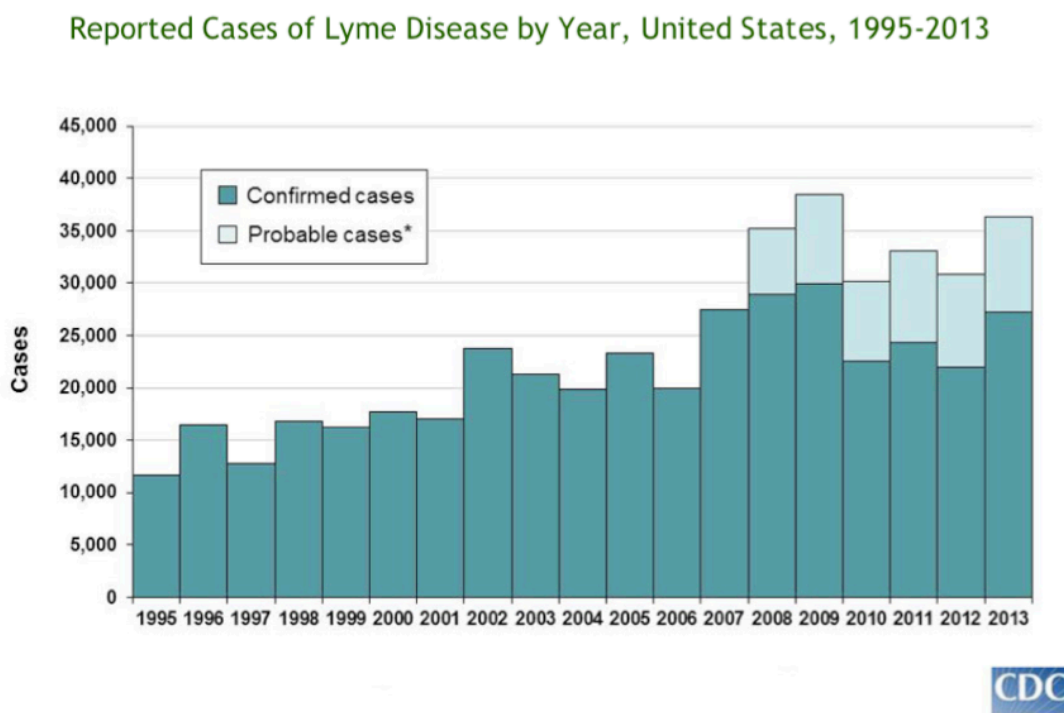


Figure 5. Cases of Lyme Disease in the United States reported to the CDC [102]. Note: From 2009-2010 the CDC enacted a change in how cases were reported.

Lyme disease presents a significant health threat, both in areas where it is endemic, and in “lower risk” areas, because of difficulties that can arise in diagnosis and recognition [103].

In addition to this health threat, an important concern remains for patients who continue to experience symptoms after the completion of what is considered to be an acceptable antibiotic regimen. This condition, commonly called Post Treatment Lyme disease Syndrome (PTLDS), can affect up to 13% of patients who have been afflicted with Lyme disease [104].

Three hypotheses [78, 88, 105, 106] have been proposed to explain PTLDS, and a combination of one or more of the following hypotheses may fully explain this phenomenon.

- 1) Since it has been shown that the OspA antigen bears close similarity to the HLA-DRB1 allele, PTLDS may be autoimmune-related [88, 106].
- 2) Attenuated *Borrelia burgdorferi*, noninfectious *Borrelial* particles, or fragments from *B. burgdorferi* could remain at target sites in the body, and continuously elicit symptoms after the cessation of antibiotic therapy [107].
- 3) A low level of infection persists in a host and continues to cause symptoms, despite immune pressure and antibiotic therapy [78, 105].

Although a combination of these hypotheses is possible, the focus of this work will be on the last hypothesis, to examine how populations of *Borrelia burgdorferi*

can alter their phenotype and adapt to survive in an unfavorable or stressful environment.

2.2 Central Hypothesis

Therefore, the central hypothesis is that *Borrelia burgdorferi* spirochetes have the ability to respond to antibiotic stress in their environment, and can adopt a slow-growing “persister” phenotype to promote survival.

2.3 Introduction and Specific Aims

Different antibiotics have diverse mechanisms of action to specifically target a component of the bacteria, and either (or both) inhibit growth, or kill the bacteria. For example, doxycycline affects the 30S ribosomal subunit to inhibit translation [108, 109], while cephalosporin antibiotics block peptidoglycan synthesis in the bacterial cell wall [110]. With the discovery of antibiotics, bacteria have evolved mechanisms of resistance to these antibiotics, such as efflux pumps [111], and beta-lactamase enzymes [112]. Bacteria have also shown the ability to spontaneously enter a state of dormancy when an antibiotic triggers a key stress-response gene, a phenomenon known as antibiotic persistence [75].

Persistence is fundamentally different from resistance. A bacterial population demonstrating resistance will continue growth in spite of the presence of an antibiotic. The mechanism of resistance will enable metabolic activity to continue, and prevent the harmful effects of the antibiotic. A mechanism of

persistence will halt metabolic activity in a subpopulation, until the antibiotic is removed. The antibiotic will kill any bacteria that are not in a persister subpopulation, but those within a persister subpopulation will survive by activating stress-response genes, and entering a state of dormancy during the antibiotic treatment. When the antibiotic is removed, the subpopulation will exit dormancy and regrow [75]. Persistence was first proposed as early as 1944 by Dr. Bigger, when he studied the effects of penicillin on *S. aureus* cultures [113], and more recently with *E. coli* and *P. aeruginosa* [114-116].

B. burgdorferi has been shown to survive both *in vitro* and *in vivo* after treatment with antibiotics [78, 105, 117]. While there have not been any clearly defined resistance mechanisms for *B. burgdorferi*, the murine model and a non-human primate model have both been utilized to study aspects of the *B. burgdorferi* infection and persistence. *B. burgdorferi* infection in the murine model can be accomplished either by tick-infection, or by needle-inoculation, and the C3H/He mouse strain has demonstrated arthritis after infection, which is a sign of Lyme disease [82, 118]. When mice were infected with *B. burgdorferi* by needle inoculation, *B. burgdorferi* DNA and RNA were detectable up to 12 weeks later in untreated, or ceftriaxone- or doxycycline-treated mice [105, 117, 119]. Infection with *B. burgdorferi* has been demonstrated in the Rhesus Macaque both with a tick vector and with needle inoculation [80, 81, 120]. The Rhesus Macaque demonstrates many signs of Lyme disease after infection with *B. burgdorferi*, including an erythema migrans and neurological pathology, and it

has been shown that after treatment with doxycycline and ceftriaxone, *B. burgdorferi* DNA and intact spirochetes are both detectable.

Several stress-response genes, and stress-response pathways have been proposed [59-64] which could enable *B. burgdorferi* to survive if it is exposed to unfavorable conditions, including antibiotic treatment. Previous work has also suggested that during treatment, *B. burgdorferi* may adopt a dormant, persister state to survive what would be a bactericidal dose of the antibiotic [105, 117, 119, 120]. These data suggest a possible mechanism of persistence for *B. burgdorferi* after antibiotic therapy. To address our central hypothesis, the following two specific aims are proposed, addressed in Chapter 1 and Chapter 2, respectively.

2.3.1 Specific Aim 1

Specific Aim 1: To determine the effect of antibiotic pressure on phenotype.

Hypothesis: B. burgdorferi spirochetes that survive in an antibiotic environment remain viable after the drug is withdrawn.

2.3.2 Specific Aim 2

Specific Aim 2: To determine the effect of bacteriostatic antibiotics on population dynamics.

Hypothesis: Observation of B. burgdorferi population changes in antibiotic-rich, and antibiotic-free media can be used to predict the probability, and explore the mechanism, of persister formation.

CHAPTER 3: EFFECT OF ANTIBIOTIC PRESSURE ON PHENOTYPE

3.1 Introduction

Lyme Borreliosis is caused by the spirochete, *Borrelia burgdorferi* (*B. burgdorferi*), and is typically treated with doxycycline, penicillin, or ceftriaxone [16]. While antibiotic therapy usually eradicates the *B. burgdorferi* infection, and results in the resolution of Lyme disease, some patients can continue to experience symptoms after the completion of antibiotic therapy. This phenomenon is known as post treatment Lyme disease syndrome (PTLDS), and there is no clear consensus in the scientific and medical communities on the cause, or the accepted treatment [106]. Three published hypotheses [78, 88, 105, 106] have been proposed to explain this phenomenon: autoimmune, noninfectious particles, and persistent infection. For a more detailed description of these hypotheses, see section 2.1 in the General introduction.

For other species of bacteria, it has been suggested [121] that persistence after antibiotic treatment is driven by the antibiotic triggering the expression of stress-response genes, which causes random phenotypic changes in subpopulations of a larger population. These phenotypically different

subpopulations would remain dormant until the removal of the antibiotic, and thus enable the survival of the bacteria. A bacterial population that grows in the presence of an established bactericidal antibiotic concentration can be described as resistant to the antibiotic. Conversely, a population which is treated with an established bactericidal antibiotic concentration, and does not grow, but then regrows after the removal of the antibiotic, can be described as persistent to the antibiotic [113]. Previous work *in vitro* [105, 119] and *in vivo* [78, 117] has demonstrated that *B. burgdorferi* spirochetes have the capability to survive antibiotic treatment. Since growth of *B. burgdorferi* during the antibiotic treatment of the *in vitro* and *in vivo* studies was not evident, this suggests some mechanism of persistence for the *B. burgdorferi* spirochete. Therefore, we propose the first Specific Aim, to determine the effect antibiotic pressure has on phenotype, and hypothesize that *B. burgdorferi* spirochetes that survive in an antibiotic environment remain viable after the drug is withdrawn.

To test this hypothesis, we proposed two sub-aims. First, we conducted an assay that increased doxycycline concentration in multiple *B. burgdorferi* population densities. The results from this assay suggested that the Minimum Bactericidal Concentration (MBC) of doxycycline was 50 $\mu\text{g}/\text{mL}$. Next, we conducted a test for viability after treatment, dubbed the “*in vivo* viability assay”. *B. burgdorferi* were treated with the Minimum Bactericidal Concentration (MBC) of doxycycline, then needle-inoculated into SCID and C3H mice to ascertain if the doxycycline treatment would prevent infection and host-adaptation. Second, an

RNA Sequencing (RNASeq) experiment was conducted to determine the expression profiles of *B. burgdorferi* before and after treatment. The RNASeq experiment was performed by extracting total RNA from untreated, treated, and treated then regrown culture-grown samples, and then processing the total RNA for sequencing with Life Technologies® PGM™. We found that after treatment with the MBC of doxycycline, *B. burgdorferi* could successfully host-adapt and infect both SCID and C3H mice. The RNASeq analysis also found differentially expressed genes, which are summarized in the results section.

3.2 Methods

***Borrelia burgdorferi*.** Low passage (p4 or p5) *Borrelia burgdorferi sensu stricto* strain B31 clonal isolate 5A19 [122] was grown in a tri-gas incubator set at 5% CO₂, 3% O₂, and the rest N₂ at 34° C in BSK-II media, as described previously [86]. The *B. burgdorferi* were seeded at low concentration from a frozen glycerol stock, and then grown to the necessary cell density.

Mice. Five to six Week old CB17 SCID and C3H/HeN C3H female mice were purchased from Charles River Laboratories. The mice were housed and maintained in filter-top cages, with water and food *ad libitum*, under the care and supervision of Tulane National Primate Research Center Veterinarians. At the beginning of the *in vivo* viability assay, pre-immune blood draws were performed by retro-orbital collection while the mice were under anesthesia. Ear-punch

biopsies and post-infection blood draws were performed at 1 week and 2 weeks post-infection, on anaesthetized mice. After 3 weeks, the mice were euthanized by carbon dioxide narcosis and cervical dislocation. A final blood draw was performed at this time. The mice ear, heart, bladder, spleen, and tibial-tarsal joint tissue was collected, and for each tissue specimen, $\frac{1}{2}$ was collected for RT-PCR, and the remaining $\frac{1}{2}$ was divided equally for PCR, and tissue culture in BSK-H. The tissues were incubated up to 30 days in a tri-gas incubator, and checked 2-3 times a week for *B. burgdorferi* growth. The Tulane National Primate Research Center (TNPRC) Institutional Animal Care and Use Committee (IACUC) approved all experiments and methods. The Tulane National Primate Research Center is a fully accredited Institution by the Association for the Assessment and Accreditation of Laboratory Animal Care International.

MBC assay. An assay to determine the minimum bactericidal concentration (MBC) for doxycycline was prepared as follows. The *B. burgdorferi* were grown in BSK-II media to stationary phase (2×10^8 cells/mL) or mid log phase (2×10^7 cells/mL). Cultures of 2×10^8 cells/mL were diluted to medium (1×10^6 cells/mL) and high (1×10^7 cells/mL) cell density, whereas cultures of 2×10^7 cells/mL were diluted to low (2×10^5 cells/mL) and medium (1×10^6 cells/mL) cell density. In all instances, the diluted cultures were treated with concentrations of doxycycline hyclate ranging from 0.1-50 $\mu\text{g/mL}$ for 5 days. A subculture of the *B. burgdorferi* was taken by removing half of the culture on day 5, gently pelleting at

3500 RPM for 20 minutes at room temperature, and re-suspending in doxycycline-free media. The cultures were checked to verify the pellet was not lost after spinning. After 5 days, the cultures were monitored for any signs of motility. The lowest concentration of antibiotic where no motility was detected at any dilution was indicated as the MBC for doxycycline with *B. burgdorferi*.

In Vivo Viability Assay. Previously, the Minimum Inhibitory Concentration (MIC) of doxycycline, or the concentration that inhibited growth, had been experimentally determined to be 0.5 $\mu\text{g/mL}$. Concentrations of 0.1 $\mu\text{g/mL}$, 0.25 $\mu\text{g/mL}$, and 0.5 $\mu\text{g/mL}$ of doxycycline were added to cultures of *B. burgdorferi*, and at 0.5 $\mu\text{g/mL}$, growth was inhibited over a 5-day interval (see Figure 6).

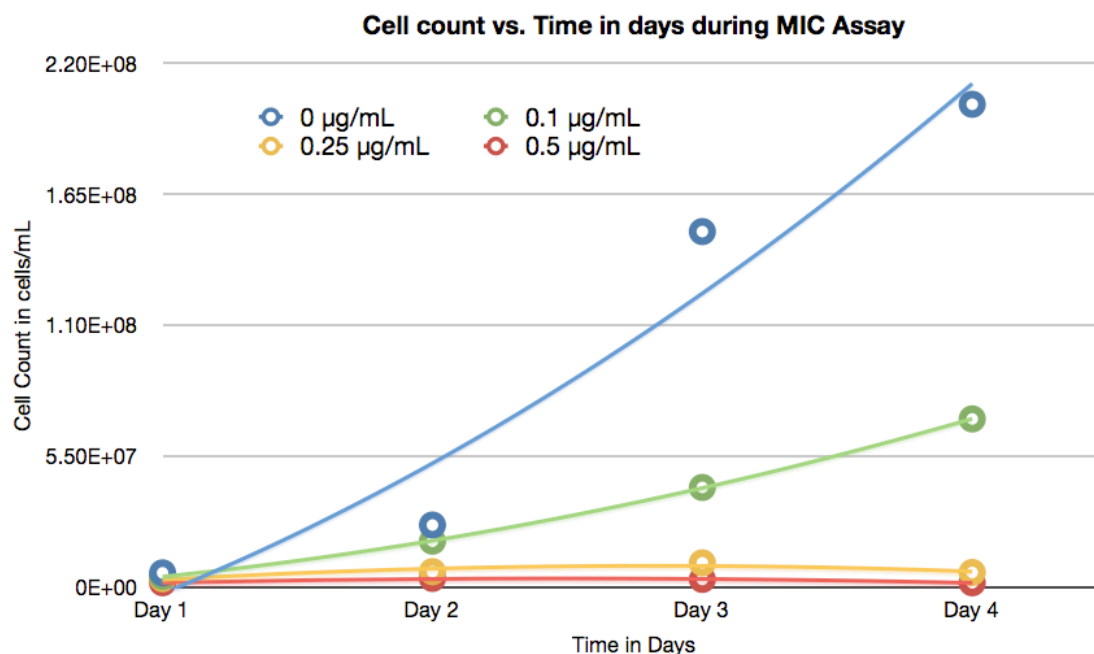


Figure 6. Cell count over 5 day doxycycline treatment. *B.b.* at 3×10^6 cells/mL concentration were treated with increasing concentrations of doxycycline to determine the minimum inhibitory concentration. A concentration of 0.5 $\mu\text{g/mL}$ successfully inhibited growth in the cultures.

Three groups of cultures of *B. burgdorferi* were grown to 5×10^7 cells/mL (See Figure 7).

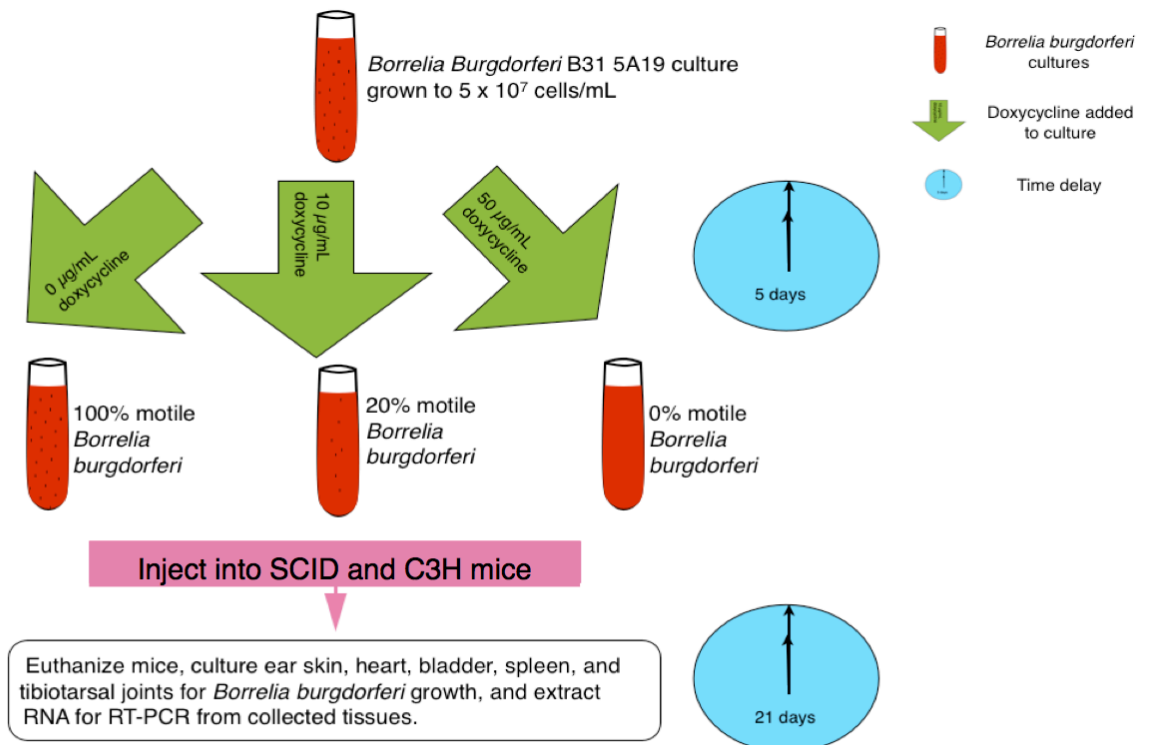


Figure 7. Overview of Experimental Process for Mouse Assay. *Borrelia burgdorferi* B31 5A19 are treated with doxycycline for 5 days, then inoculated into SCID or C3H mice, then tissues are collected after the mice are euthanized.

Next, the cultures were either not treated, treated with a concentration higher than the MIC but not greater than the MBC (10 µg/mL doxycycline), or treated with the MBC (50 µg/mL doxycycline). At day 5, the cultures were checked for motility. The untreated control had 100% motile *B. burgdorferi* in the culture, the culture treated with 10 µg/mL had approximately 20% motile *B. burgdorferi* in the culture, and the culture treated with 50 µg/mL had 0% motile *B. burgdorferi* in the culture. A sample of 2×10^5 *B. burgdorferi* was needle-inoculated into SCID mice and C3H mice, after a pre-immune blood-draw had been taken from the C3H mice. Blood draws were taken from the C3H mice on day 7 and day 14, and ear-

punch biopsies were taken on day 14 for all mice. On day 21, the mice were euthanized, and tissues were harvested for tissue culture in BSK-H in 5 mL snap-cap tubes, and in RNALater™ to perform RT-PCR and probe for both *flaB* and *ospC*. The tubes were incubated in a tri-gas incubator as described above, and checked 2-3 times weekly, for up to 30 days, for signs of motility. The experiment was repeated three times, with a total of five control group SCID mice and C3H mice, a total of three 10 µg/mL doxycycline-treated-group SCID mice and C3H mice, and a total of seven 50 µg/mL-doxycycline-treated-group SCID mice and C3H mice.

BacLight Staining. To determine the proportion of live and dead *B. burgdorferi* following doxycycline treatment, Live/Dead BacLight® (Molecular Probes) staining was performed on untreated and 50 µg/mL-treated *B. burgdorferi* per the manufacturer's instructions. Briefly, 1.0 µL of 1.67 mM SYTO-9, a green nucleic acid stain, and 1.0 µL of 1.67 mM propidium iodide, a stain which will only stain cells with damaged cell membranes red, were thawed and mixed in equal proportions. 1.5 µL of the stain mixture was then added to 500 µL of *B. burgdorferi* suspended in PBS, incubated in the dark for 15 minutes, then applied to a slide and coverslip for viewing and counting. The samples were viewed and counted in live/dead ratios using a Leica Fluorescent Microscope. The excitation/emission maxima for Syto9 is 480/500 nm, and 490/635 nm for propidium iodide. Images captured were obtained using a Nuance FX®

fluorescence microscope and software, with an optimal emission filter range of 500-560 nm for Syto9, and 600-650 nm for propidium iodide. Results are calculated as percent viability, and reported as mean +/- SD per group.

RT-PCR. During the *in vivo* viability assay, mice tissues were harvested and placed into RNALater until processing. During processing, the tissues were ground with a tissue grinder, then the RNA was extracted using a Qiagen® RNeasy™ Kit and the RNeasy™ Fibrous Tissue Kit. The tissues were probed for *flaB* (flagellin, used in identification, as described in [123]), *oppA2* (conserved peptide epitope, with sequence *oppA2* REV 5'-GAG CCT CGA GTT ATT TAT TTT TTA ATT TTA GCT G -3', see also [124]), and *ospC* (Outer Surface Protein C, associated with virulence and host-infection, as described in [123]).

RNASeq. A differential expression experiment to determine gene expression under antibiotic-treated, and antibiotic-withdrawn conditions was conducted, with three groups and one replicate of each group. Group 1, B31, contained RNA from an untreated *B. burgdorferi* control. Group 2, B31P, contained RNA from *B. burgdorferi* treated with 50 µg/mL doxycycline for 5 days. Group 3, B31PG, contained RNA from *B. burgdorferi* treated with 50 µg/mL doxycycline for 5 days, then allowed to regrow until the population reached the initial concentration of B31PG before treatment.

Total RNA was extracted and processed for sequencing following methods outlined in Appendix B. After sequencing, the reads were aligned to the B31 5A19 RefSeq genome at the NCBI database, Accession number PRJNA57581 [125], using Ion Torrent® Suite Server. Post-processing analysis was performed using Partek® Genomics Suite® Software, ©2013, using a statistical significance of $p < 0.05$ calculated using the log-likelihood test in Partek® Genomic Suite™ Software, ©2013. The full results of the analysis are available at the online database <http://borrelia.rna.technology> (See Appendix C) and a detailed description of RNASeq theory and methods used can be found in Appendix B. An overview of the calculations performed can be found in section 1.2 of Appendix A.

Statistical Analysis. Analysis of Variance (ANOVA), Student's t-test, Chi-Squared, and Fisher's exact test were performed with GraphPad® Prism™ and with R (<http://www.R-project.org>). RNASeq statistics were performed using Partek® Genomic Suite™ Software, ©2013.

3.3 Results

We reasoned that although doxycycline was bacteriostatic, it acted on the 30S ribosomal subunit [126] to inhibit protein synthesis, so it could effectively eliminate a bacterial population if the concentration and duration were sufficient. Therefore, an MBC assay with differing population densities was set up to

determine the concentration of doxycycline that would prevent growth and motility after 5 days. After 5 days, the MBC for the sub-cultures was determined by quantifying the number of motile spirochetes (See Figure 8).

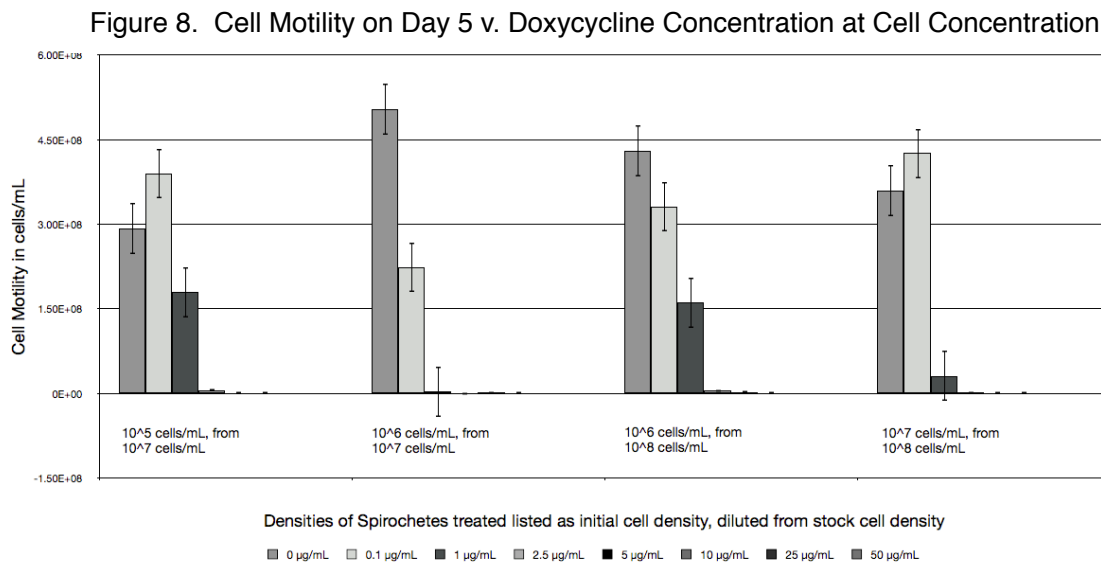


Figure 8. Cell Motility on Day 5 of MBC Assay. A growth inhibition assay was performed, whereby motility in cultures was measured on day 5 with differing initial concentrations and differing seed concentrations. The results are displayed in log-transformed results for readability. No motility was observed at any cell density at 50 µg/mL doxycycline concentration.

There was no motility observed on day 5 for any culture or cell density that was treated at 50 µg/mL, suggesting an MBC value of 50 µg/mL for doxycycline.

To determine the effectiveness of the treatment, in addition to motility, BacLight staining was performed on *B. burgdorferi* that were untreated, and those treated with 50 µg/mL of doxycycline (see Table 2, Figure 9).

Table 2. BacLight Staining of Untreated and Treated <i>B. bugrdorferi</i>			
	Live	Dead	%Live
Control	155 +/- 18	12 +/- 9	92.8%
Treated	21 +/- 4	37 +/- 15	36.2 %
Significance	p < 2.2 x 10 ⁻¹⁶ , Fisher's Exact Test		

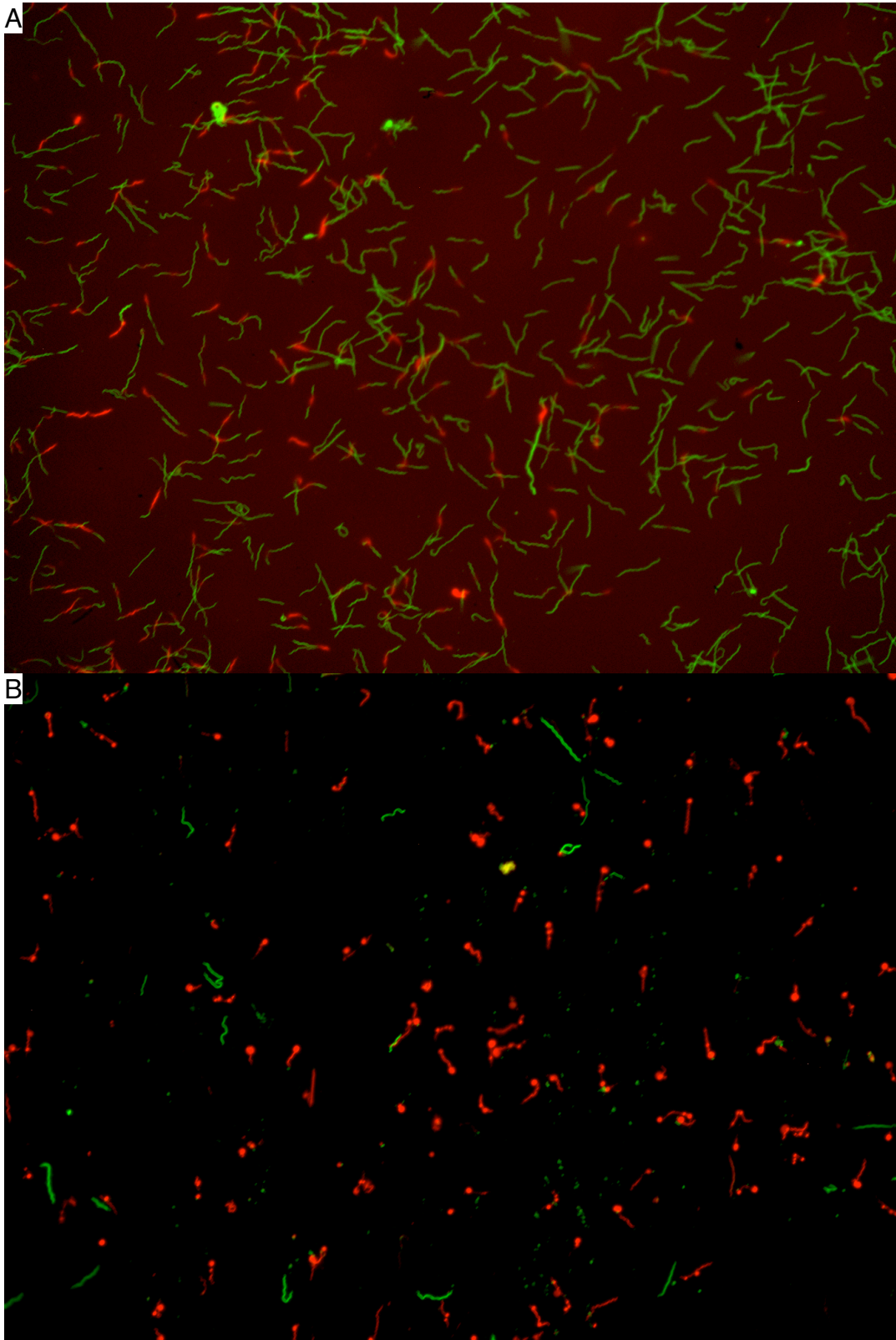


Figure 9. BacLight Image of *B. burgdorferi*. *B. burgdorferi* were either (A) not treated with doxycycline or (B) treated with 50 $\mu\text{g}/\text{mL}$ of doxycycline for five days. On day 5, the cultures were stained with a BacLight stain and observed with fluorescence microscopy. Green are live, red are dead.

The results suggest that the MBC for doxycycline had a significant effect on membrane integrity and viability.

We initially hypothesized that during treatment with the MBC of doxycycline, a subpopulation would adopt a persister phenotype. Then, if the treated *B. burgdorferi* were needle-inoculated into SCID mice, the subpopulation could host-adapt and successfully infect the host, but if the *B. burgdorferi* were needle-inoculated into immune-competent C3H mice, host immune pressure would prevent infection. The treatment had a significant effect on membrane integrity and viability. As can be seen in Table 3, the results suggested that the first hypothesis was correct, but the second hypothesis was incorrect.

Table 3a. Results with SCID mice, 3 experiments		
<i>B.b.</i> Treatment	Motility of inoculation <i>B.b.</i>	Culture Positive*
0 $\mu\text{g}/\text{mL}$ doxycycline	100%	4/5
10 $\mu\text{g}/\text{mL}$ doxycycline	10-20%	3/3
50 $\mu\text{g}/\text{mL}$ doxycycline	0%	1/7
*number positive/total mice tested		

Table 3b. Results with C3H mice, 2 experiments		
<i>B.b.</i> Treatment	Motility of inoculation <i>B.b.</i>	Culture Positive*
0 $\mu\text{g}/\text{mL}$ doxycycline	100%	3/5
10 $\mu\text{g}/\text{mL}$ doxycycline	10-20%	1/3
50 $\mu\text{g}/\text{mL}$ doxycycline	0%	2/7
*number positive/total mice tested		

The treated *B. burgdorferi* were able to host-adapt after treatment, and successfully infect SCID and C3H mice. When sample RNA from untreated spirochetes, 10 $\mu\text{g}/\text{mL}$ spirochetes, and 50 $\mu\text{g}/\text{mL}$ spirochetes was probed with *flaB* and *ospC* (see Figure 10), the 10 $\mu\text{g}/\text{mL}$ treated samples which infected both the SCID mice and the C3H mice both were positive.

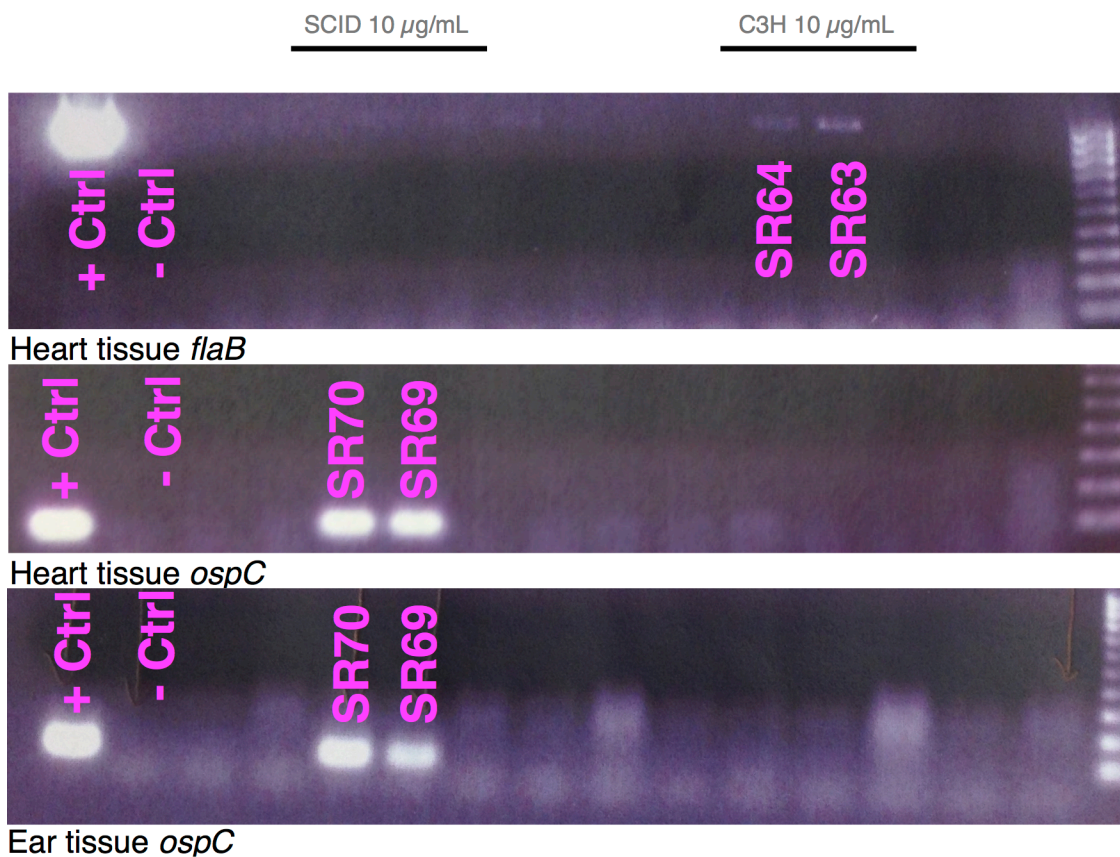


Figure 10. Mouse RT-PCR. From Top to Bottom, Heart tissue RT-PCR amplified with primers for *flaB*, heart tissue RT-PCR amplified with primers for *ospC*, and ear tissue RT-PCR amplified with primers for *ospC*. C3H mice SR63 and 64, and SCID mice 69 and 70 are positive.

For reasons that are unclear, particularly since motility was observed in tissue culture for these groups, RT-PCR results were either negative, or inconclusive, for mice that were infected by the untreated- and the 50 $\mu\text{g}/\text{mL}$ -treated *B.*

burgdorferi spirochetes.

For the RNASeq analysis, we set up the experiment as described in Appendix B, and in the Methods section. Briefly, total RNA from untreated, treated, and treated then regrown *B. burgdorferi* was extracted, depleted of ribosomal RNA, fragmented and converted into a cDNA library template, then submitted for sequencing with the Life Technologies Ion Torrent® PGM™. The post sequencing analysis was conducted using Partek® Genomics Suite® to determine genes which were differentially expressed between treated and untreated *B. burgdorferi* samples. We hypothesized that if RNA from a control sample of *B. burgdorferi* were compared to RNA from a treated sample of *B. burgdorferi*, and likewise if a control sample were compared to a sample which had been treated and then allowed to regrow, then genes which were differentially expressed before and after doxycycline treatment could be determined. Although the experiment lacked replicates, Partek® Genomics Suite® software was able to calculate differential expression for gene sets between the untreated and treated, and the treated and treated/regrown groups.

The data revealed that genes in several plasmids, like linear plasmid 5 (lp5) and circular plasmid 9 (cp9), were upregulated in the RNA from the treated sample compared to RNA from the control sample, but this trend was reversed in genes from the treated and regrown sample compared to genes from the untreated sample (See Figure 11).

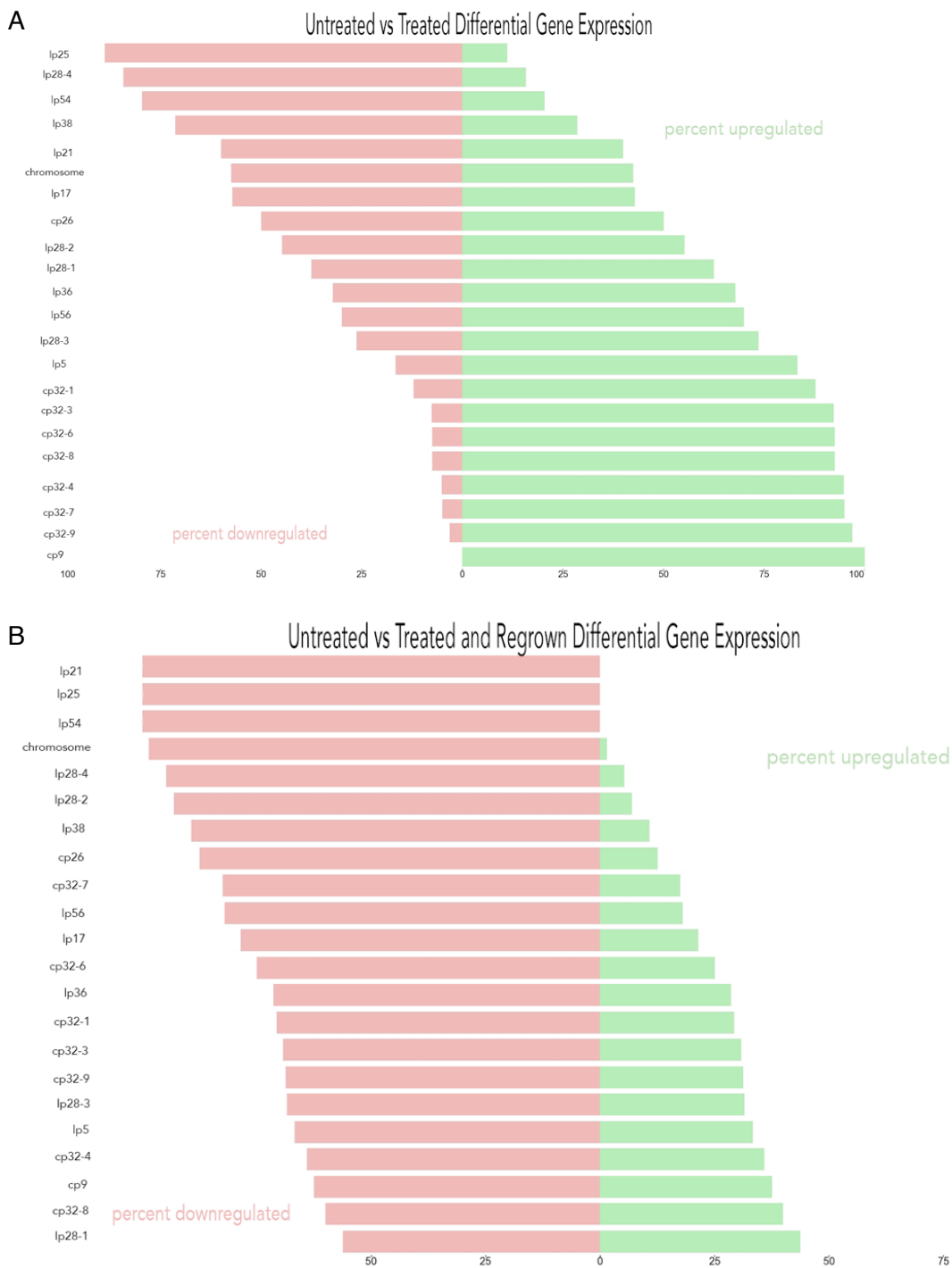


Figure 11. Forest Plot of Gene Upregulation and Downregulation in RNASeq Data
 A comparative analysis from the *B. burgdorferi* RNASeq data, showing which genomic elements may be upregulated or downregulated (A) in the treated compared to the untreated sample, and (B) which genomic elements may be upregulated or downregulated in the treated and regrown sample compared to the untreated sample.

Interestingly, nearly all genes in the chromosome show downregulation in the

treated and regrown sample compared to the control sample, while in the treated sample, more genes were upregulated (Figure 11). A Venn Diagram of 2-fold change or greater in expression shows 178 genes that were expressed, and genes expressed under treatment or regrowth may be stress-response pathway genes (Figure 12).

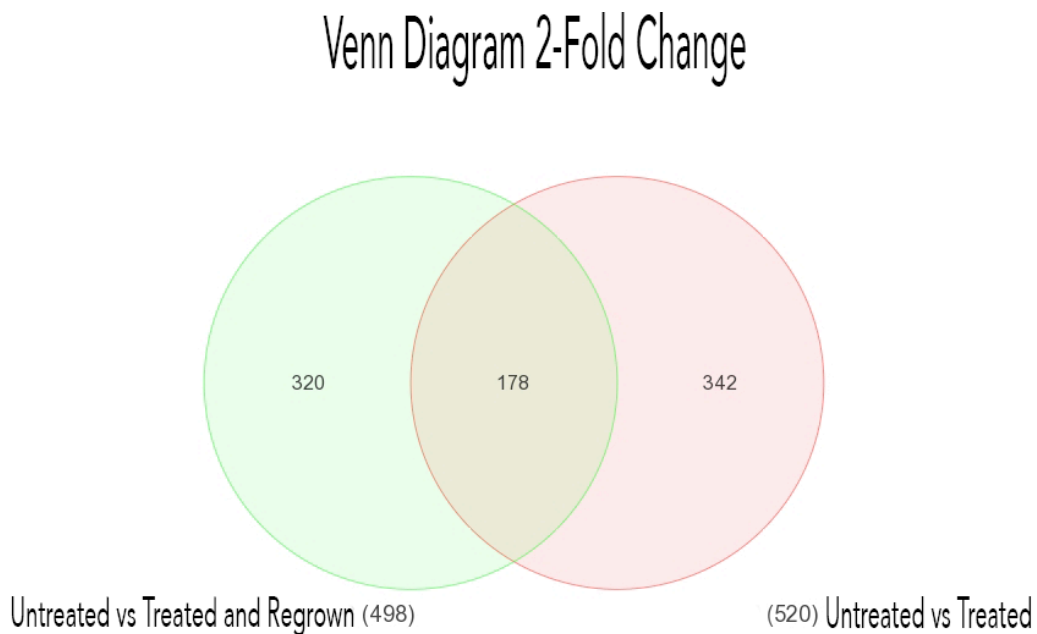
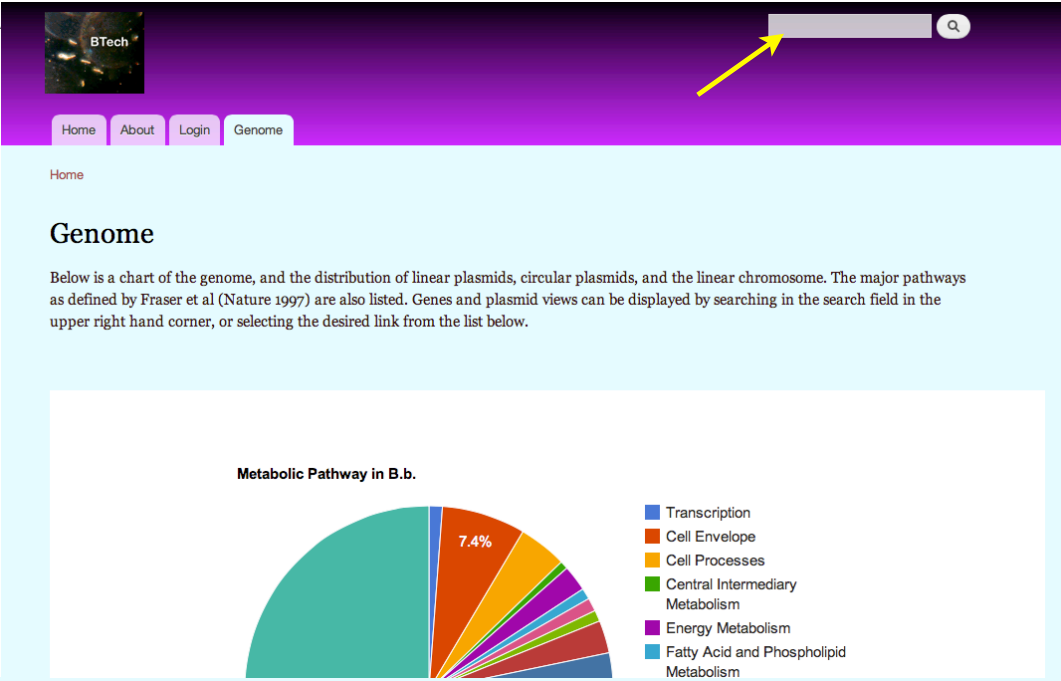


Figure 12. Venn Diagram Showing genes that are significantly expressed above the 2-fold level

An abbreviated dataset is presented in Table 4 from the Partek analysis, based on genes which are significantly different between samples and may be 2-fold upregulated. The full dataset is available at the database that was created at <http://borrelia.rna.technology> (See Appendix C). The genome of *B. burgdorferi*

has been annotated with relevant genes, and the results of the RNASeq analysis have been uploaded to the server. The results are accessible by the search field, genome page, and a pathway page (See Figure 13).

A



Home About Login Genome

Home

Genome

Below is a chart of the genome, and the distribution of linear plasmids, circular plasmids, and the linear chromosome. The major pathways as defined by Fraser et al (Nature 1997) are also listed. Genes and plasmid views can be displayed by searching in the search field in the upper right hand corner, or selecting the desired link from the list below.

Metabolic Pathway in B.b.

- Transcription
- Cell Envelope
- Cell Processes
- Central Intermediary Metabolism
- Energy Metabolism
- Fatty Acid and Phospholipid Metabolism

B

Name	Genomic Type	Notes or Notable Genes/Proteins
lp5	linear plasmid	Outer Membrane Protein, hypothetical proteins
lp17	linear plasmid	protein CdsM, other pseudogenes
lp21	linear plasmid	predicted proteins
lp25	linear plasmid	protein p23, required for infectivity[1,2]
lp28-1	linear plasmid	transmembrane protein, required for infectivity[1,2]
lp28-2	linear plasmid	DNA helicase, predicted proteins
lp28-3	linear plasmid	predicted proteins
lp28-4	linear plasmid	virulence-associated lipoprotein
lp36	linear plasmid	fibronectin-binding protein, predicted proteins
lp38	linear plasmid	opp proteins, ospD
lp54	linear plasmid	outer membrane proteins, ospA, ospB
lp56	linear plasmid	ErpX, BppA
cp26	circular plasmid	ospC, opp proteins
cp32-3	circular plasmid	BppB proteins
cp32-4	circular plasmid	predicted proteins
cp32-6	circular plasmid	predicted proteins
cp32-7	circular plasmid	predicted proteins
cp32-8	circular plasmid	predicted proteins
cp32-9	circular plasmid	Erp-family, Bpp-family proteins, predicted proteins
cp9	circular plasmid	predicted proteins
chromosome	cchromosome	tRNA, rRNA, metabolic proteins

[1] Purser and Norris. "Correlation between plasmid content and infectivity in *Borrelia burgdorferi*". PNAS 2000.
 [2] Labandeira-Rey and Skare. "Decreased Infectivity in *Borrelia burgdorferi* Strain B31 Is Associated with Loss of Linear Plasmid 25 or 28-

Figure 13. BTech: Online, Searchable Database for *Borrelia burgdorferi* genome and transcriptome. The database is accessible at <http://borrelia.rna.technology> (A) The yellow arrow shows a search field where searches can be entered, and (B) the yellow arrow shows the column where links can be clicked to navigate to genes that are listed on that plasmid or chromosome.

The pathway page is a specific webpage on the server that been set up with genes that are differentially expressed from the BosR *B. burgdorferi* pathway.

Table 4. Selection of Genes Found to be Significantly Differently Expressed After Treatment with Doxycycline	
Gene Name Or Description	Gene ID
30 S ribosomal protein subunit-encoding genes	multiple including BB_0114, BB_0127
50 S ribosomal subunit gene	BB_0497
50 S ribosomal general stress protein	BB_0786
transcription cleavage factor	BB_0132
Rho transcr term factor	BB_0230
cell division protein-encoding genes	BB_0299, BB_0300
transcription antitermination factor	BB_0394
response regulator	BB_0419
Chaperone protein-encoding gene	BB_0517
ErpX	BB_Q47
Decorin binding protein (dbpA, dbpB)	BB_A24, BB_A25
BppA, BppB	BB_Q43, BB_A23
gene translating and ATP binding protein	BB_J26
gene translating immunogenic protein P37	BB_K45

Based on the analysis from Partek, it is possible that the genes listed in Table 4 were 2-fold upregulated in the treated sample compared to the untreated sample, including a ribosomal stress-response gene, and a transcription cleavage factor.

It is also of particular interest that genes to translate ATP binding proteins were induced, which could suggest a DNA repair mechanism [127] or a dormancy mechanism. It is important to note that numerous genes were downregulated as well as upregulated, however, the current analysis limited the extent of how downregulation could be analyzed. Future RNASeq experiments will allow a broader and more in-depth analysis.

3.4 Discussion

The data demonstrated that *B. burgdorferi* can host-adapt following doxycycline treatment to establish an infection in both an immune-deficient and immune-competent host. This further suggests that after treatment with doxycycline, *B. burgdorferi* can induce expression in stress-response genes, and adopt a dormant, persister state. It has been well established that *B. burgdorferi*, can alter phenotype during its life cycle [34, 38, 54-56, 128], such as downregulating *ospA* and upregulating *ospC* shortly after infection of a mammalian host, and altering surface expression of the *vlsE* genes [70, 71] to avoid recognition by the host immune system. If *B. burgdorferi* encounters doxycycline, it may also be able to survive antibiotic therapy by inducing a key stress-response gene or pathway, and then adopting a dormant state until the removal of the antibiotic.

To answer the question of which genes are induced after treatment, we conducted the RNASeq experiment. Preliminary data from this experiment

demonstrated that several genes, including a stress-response ribosomal gene at BB_0786, were induced after treatment. These results will need to be followed up with additional analysis and biological replicates, but this opens exciting possibilities for the exploration of the *B. burgdorferi* transcriptome and genome.

CHAPTER 4: EFFECT OF BACTERIOSTATIC ANTIBIOTICS ON POPULATION DYNAMICS

4.1 Introduction

When presented with an external stimulus such as food or toxins, bacterial populations will form heterogeneous subpopulations, depending on how each subpopulation upregulates or downregulates its target genes as a response [73, 74]. As a simplified example, consider the addition of ciprofloxacin to media with *E. coli*. Ciprofloxacin inhibits bacterial topoisomerase [115] in *E. coli* cells, and as a result causes *E. coli* to induce the SOS repair response pathway [129]. Unless the bacteria happen to be resistant to ciprofloxacin by accident via a mutation, or by design via a plasmid, the induction of the SOS response will still lead to cell death [115]. However, when the SOS repair pathway is induced, it can also rarely activate *tisB*, which acts as a cellular toxin. The *tisB* gene is normally repressed, but when induced, it decreases proton motive force and metabolic activity [115, 121]. As part of a larger population, if ciprofloxacin is present, the majority of cells which attempt SOS repair will be killed, but a small

subpopulation of cells that enter a state of dormancy by inducing *tisB* until ciprofloxacin is removed will survive the antibiotic treatment [121].

It has been demonstrated in previous work [78] and in data presented in Chapter 1, that a subpopulation of *B. burgdorferi* can adopt a persister state if it is confronted with the antibiotic doxycycline. The mechanisms involved in the transition to a persister state are poorly understood, and the frequency of such an occurrence is even less understood. Since PTLDS is observed in up to 13% [104] of patients who have had Lyme disease, these observations and data suggest that similar to other bacteria that demonstrate persistence, if treated with doxycycline, some underlying randomness drives the transition to dormancy.

Quantifying the existence and emergence of a persister subpopulation within a larger population could elucidate the role of *B. burgdorferi* in PTLDS, and provide valuable information on how subpopulations alter phenotype. Therefore, our second specific aim was to determine the effect bacteriostatic antibiotics have on population dynamics. We hypothesize that the observation of *B. burgdorferi* in antibiotic-rich and antibiotic-free media could be used to predict the probability, and explore the mechanism, of persister formation.

To demonstrate this hypothesis, we proposed two sub-aims. First, to determine how growth phase affects the development of persister cells in the presence of doxycycline, we conducted a Probability Assay. This assay quantified the chance in a population that after treatment, a persister subpopulation would be present. This assay also provided data for a

mathematical model, whereby data from the Probability Assay, and a Pulse Dose Assay, modeled after experiments from Dr. Kim Lewis [75], allowed the prediction within a given population of the number of cells in a given subpopulation which adopted a persister phenotype.

4.2 Methods

Borrelia burgdorferi. Low passage (p4 or p5) *Borrelia burgdorferi sensu stricto* strain B31 clonal isolate 5A19 [122] was grown in a tri-gas incubator set at 5% CO₂, 3% O₂, and the rest N₂ at 34° C in BSK-II media, as described previously [86]. The *B. burgdorferi* were seeded at low concentration from a frozen glycerol stock, and then grown to the necessary cell density.

Probability Assay. To quantify the amount of *B. burgdorferi* that would persist after treatment with the MBC of doxycycline, an experiment was set up as follows. First, cultures of *B. burgdorferi* were grown to 2-3 x 10⁸ cells/mL, then diluted to concentrations of early log (2 x 10⁶ cells/mL), mid log (2 x 10⁷ cells/mL), and late log (1 x 10⁸ cells/mL) in 15 mL conical tubes. The early, mid, and late log groups were then treated with the MBC of doxycycline for 5 days. On day 5, the *B. burgdorferi* cultures were resuspended in doxycycline-free media. Prior to the start of the assay, it was experimentally determined that growth was most frequently observed by day 11, and if motility was not observed in the tubes by

day 11, it was unlikely to be observed at a later date (data not shown).

Therefore, the end time point was established at 11 days, and on day 11, each tube was checked for motility. Tubes with motile spirochetes were recorded as (+), and tubes with no motile spirochetes were recorded as (-).

Next, the assay was repeated by seeding the *B. burgdorferi* from a glycerol stock in 15 mL conical tubes, and growing the cultures to early log (2×10^6 cells/mL), mid log (2×10^7 cells/mL), and late log (1×10^8 cells/mL). The assay was repeated as described previously, and on day 11, the tubes with motile spirochetes were recorded as (+), while the tubes with no motile spirochetes were recorded as (-).

Pulse Dose Assay. Cultures of *B. burgdorferi* were grown to 3×10^7 cells/mL in 5 mL snap-cap tubes, treated with $50 \mu\text{g/mL}$ of doxycycline for 5 days, then pelleted and resuspended in doxycycline-free media. The culture tubes were incubated in a tri-gas incubator as described above, and monitored each day for growth. When the culture reached early-log phase ($\approx 5 \times 10^6$ cells/mL), $50 \mu\text{g/mL}$ of doxycycline was added before the *B. burgdorferi* cell density could grow to the initial assay concentration (3×10^7 cells/mL). After 5 days of doxycycline treatment, the doxycycline was removed, and the *B. burgdorferi* were monitored again for regrowth. The cycle of treating, removing the doxycycline, and monitoring for regrowth was repeated 3 times, similar to that described by Lewis *et al* [75]. Thirty-four samples were used in the pulse dose assays, and

approximately 30% of the total number of samples regrew after the first treatment. Successive treatments had a regrowth rate of approximately 50% (additional cultures were lost due to contamination). The assay was considered fully complete after three treatment cycles (three treatments and regrowth phases), and the experiment was repeated once. At the pulse dose assay endpoint, a glycerol stock was made from the *B. burgdorferi* culture for analysis and comparison with the Probability Assay. See Figure 14 for a representative diagram.

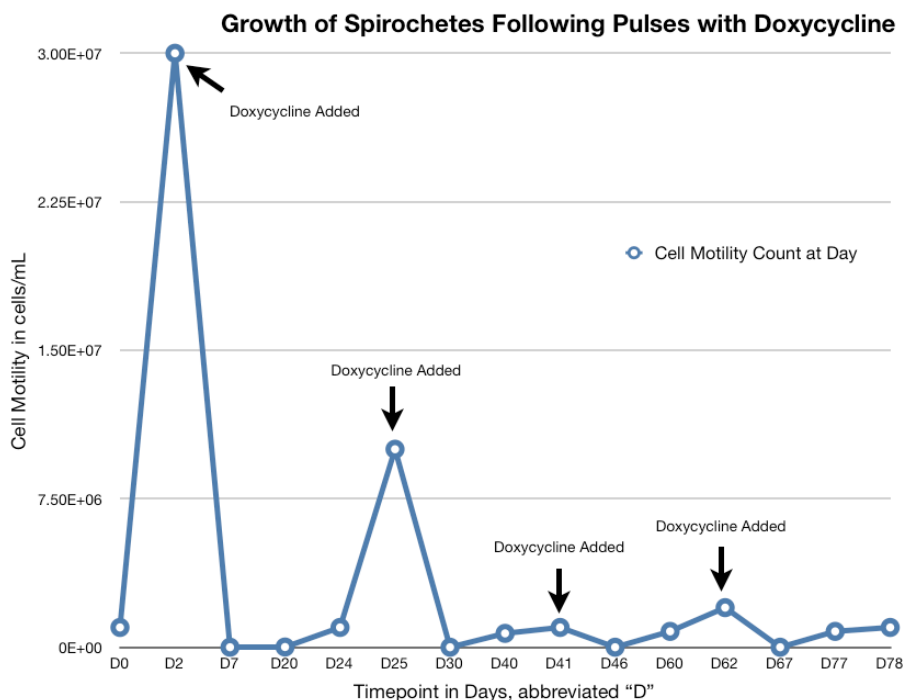


Figure 14. Representative Figure of Growth of Spirochetes Following Pulses with Doxycycline

Mathematical Modeling. To determine a mathematical model for the presence of persister subpopulations in *B. burgdorferi*, data from the Probability Assay and Pulse Dose Assay were analyzed. In a given assay with *B. burgdorferi*, for

calculation purposes, each tube was considered a population. The beginning of an assay was considered to be $t = 0$. After doxycycline was removed, the cultures with motile *B. burgdorferi* were counted on day 11 ($t = 11$) to determine the number of populations with subpopulations which transitioned from one state to another.

By examining the Pulse Dose Assay and the Probability Assay, the experimentally determined probability for regrowth after treatment during an assay of time t , $P(x) = \Sigma (\text{live})/(\text{dead})$ can be determined. Therefore, the amount of persisters $P_{\text{post-treatment}}$ in a population P during an interval of time t is defined as:

$$P_{\text{regrowth}} = P_{\text{post-treatment}} e^{kt}$$

$t = 2$ days if population size $> 1.0 \times 10^8$ cells/mL

$t = 4$ days if population size $< 1.0 \times 10^8$ cells/mL

For a detailed description and proof, see Section 1.1 of Appendix A.

Statistical Analysis. Analysis of Variance (ANOVA), student's t-test, and Fisher's exact test were performed with GraphPad Prism® and with R (<http://www.R-project.org>). The distribution error for Figure 15 was calculated as Gaussian distribution.

4.3 Results

We conducted a probability assay using the previously determined Minimum Bactericidal Concentration (MBC) for doxycycline (See Chapter 3, Figure 8), to determine whether regrowth would occur after a set time in early, mid, and stationary phase populations. The assay was performed both by diluting the cultures from a higher density to lower concentrations of early, mid, and early stationary phase, and by growing the cultures from a glycerol stock to early, mid and stationary phase. The results indicated that if the concentration was grown to stationary phase, and then diluted to a lower density, there was a significant difference in the number of cultures with motile *B. burgdorferi* in early, mid, and stationary phase groups (Table 5a).

Table 5a. Probability Assay I			
	Early	Mid	Late
Tubes with Motile Cultures on Day 11	23	36	36
Tubes with Nonmotile Cultures on Day 11	37	14	0
Tubes with Motile Cultures/ Total Tubes	23/60	36/50	36/36
Percent Tubes with Motile Cultures	38%	72%	100%
Statistical Significance (Chi-Squared with Yates' Continuity Correction)	Early v. Mid $p = 0.000857$	Mid v. Late $p = 0.00154$	Early v. Late $p = 6.89 \times 10^{-9}$

If the *B. burgdorferi* was grown to early, mid, and stationary phase, there was no regrowth or motility in any group except the stationary phase (Table 5b).

Table 5b. Probability Assay II			
	Early	Mid	Late
Tubes with Motile Cultures on Day 11	0	0	12
Tubes with Motile Cultures on Day 11	22	22	31
Tubes with Motile Cultures/Total Tubes with Cultures	22	22	43
Percent Tubes with Motile Cultures	0%	0%	28%
Statistical Significance (Chi Squared Test with Yates' Continuity Correction)	Early v. Mid null	Mid v. Late (approximation) $p = 0.0162$	Early v. Late (approximation) $p = 0.0162$

If the results of the two assays are compared together, there is a significant difference between the number of motile *B. burgdorferi* cultures in the cultures that were diluted from stationary phase to a specific population density, and the cultures that were grown to the same population density (see Table 5c).

Table 5c. Comparison of Tubes with Cultures Grown to Population Density and Tubes with Cultures diluted from Stationary Phase						
	Grown To Early Log	Diluted to Early Log From Stationary Phase	Grown To Mid Log	Diluted to Mid Log From Stationary Phase	Grown To Late Log	Diluted to Late Log From Stationary Phase
Motile on Day 11	0	23	0	36	12	36
Nonmotile on Day 11	22	37	22	14	31	0
Total	22	60	22	50	43	36
Percent	0%	38%	0%	72%	28%	100%
Statistical Significance (Pearson's ChiSq Test)	$p = 0.0016$		$p = 7.75 \times 10^{-8}$		$p = 2.89 \times 10^{-10}$	

These results suggest that population density affected the ability of *B. burgdorferi* to form persister cells within a population.

To explore the possibility that a small number of persister cells survive post-treatment, a pulse-dose assay similar to the model used by Dr. Kim Lewis [75, 121, 130] with *E. coli* and *P. aeruginosa* was established with the Pulse Dose Assay experiment. We expected that the persister population would be reduced by each subsequent treatment, and would be evidenced by lack of regrowth, or slower regrowth. Briefly, *B. burgdorferi* were grown to 2×10^8 cells/mL, diluted to a lower density, and then allowed to grow until the concentration reached 3×10^7 cells/mL. At this point, the culture was treated with the MBC ($50 \mu\text{g/mL}$) of doxycycline for 5 days. The doxycycline was then removed on day 5, and the culture was monitored for regrowth. If the population

regrew to early log phase (2×10^6 cells/mL), the process was repeated before the cell density reached the concentration before treatment. The treatment was unable to kill the persisters, which regrew after each treatment (Figure 14). Furthermore, the rate of regrowth did not significantly decline (Figure 15).

Figure 15. Average Time to Regrowth during Pulse Dose Assay for *B. burgdorferi* after End of Doxycycline Treatment

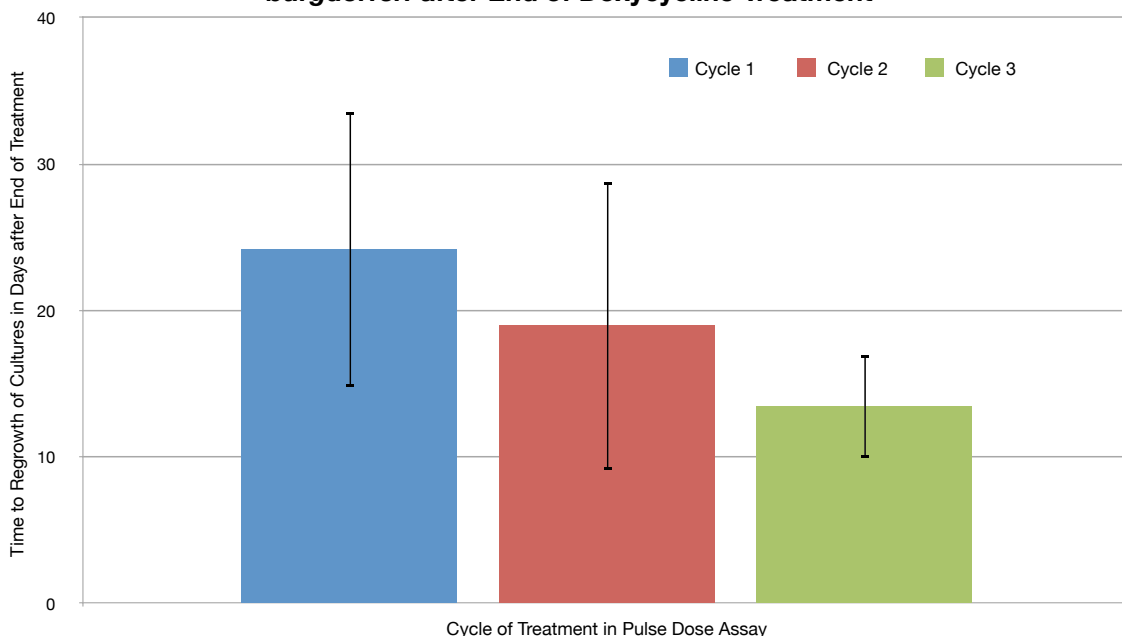


Figure 15. Pulse Dose Assay Time to Regrowth. (A) Representative figure of the Pulse Dose Assay. (B) Time Growth Measurements. *B. burgdorferi* cultures ($n = 34$) were seeded initially at 7×10^5 cells/mL, treated with the MBC of doxycycline, then monitored for regrowth, and the cultures which regrew ($n = 5$) were treated before they reached their initial concentration. The cultures were then treated for 5 days with the MBC of doxycycline, and the process was repeated two more times for the cultures which regrew ($n = 2$). Figure displays the time to regrowth after cycles 1, 2, and 3 of treatment with doxycycline and regrowth. Error bars are SD. No significant difference was observed in the time to regrowth between Cycle 2 and 3.

Finally, we hypothesized that repeated administration of antibiotics (or pulses) could create a single culture of *B. burgdorferi* persisters. To test this hypothesis, we used an isolate from the pulse dose assay with the probability assay (Figure 14). Although *B. burgdorferi* did regrow after treatment, there was a significant difference between the B31 5A19 isolate and the pulse dose assay isolate (See Table 6).

Table 6. Comparison of Tubes with Motile Cultures Between Pulse Dose Assay Isolate and B315A19 Isolate

	Early Log B31 5A19	Early Log Persister Isolate	Mid Log B31 5A19	Mid Log Persister Isolate	Late Log B31 5A19	Late Log Persister Isolate
Tubes with Motile Cultures	23	5	36	10	36	9
Tubes with Nonmotile Cultures	37	26	14	20	0	17
Tubes with Motile Cultures/Total Tubes	23/60	5/31	36/50	10/30	36/36	9/26
Percent Motile	38%	16%	72%	30%	100%	34%
Statistical Significance	p = 0.0355 (Fisher exact test)		p = 0.001 (Fisher exact test)		p = 4.224 x 10 ⁻⁹ (Fisher exact test)	

This suggests that adoption of the persister phenotype in *B. bugdorferi* is a stochastic phenomenon, and although it is more likely to occur at stationary phase, it is governed by random events.

If the results of the Probability Assay and the Pulse Dose Assay are analyzed together, several useful data points can be obtained. As cell density increases and progresses to stationary phase, there is an increasing probability for a bacterial subpopulation to alter its phenotype to a persister state, and similarly, the time required for regrowth by the subpopulation will decrease. At earlier growth phases, a persister population may exist, but the time *t* to regrowth is much greater, and the chance of regrowth is much less than 50% (See Table

5). At stationary phase, the time t that is required for regrowth is much less, and the chance of regrowth is increased compared to populations at a density less than stationary phase (see Table 5 and Table 6).

To create a predictive model, which could ascertain in any given population, the number of persisters in a population of *B. burgdorferi*, we must first assume that growth rate is constant within the cycle of regrowth and treatment, and if growth occurs, it will be exponential and follow the exponential growth rate

$$P = P_0 e^{kt} ; \text{ where}$$

P = final population concentration

P_0 = initial population concentration

$e = e$

k = growth rate constant

t = time for population to grow from P_0 to P

If a timeline of growth, addition of doxycycline, removal of doxycycline, then regrowth is visualized, then at some point after the removal of doxycycline but before motility is observed, regrowth occurs at a time $t_{regrowth}$. The value of the k constant can be derived by examining the count of the population when motility was first observed, and the count of the population before the next addition of doxycycline. Quantifying the initial concentration of persisters, and

therefore the concentration of a persister population in the population at the start of the cycle is a bit problematic, since neither $t_{regrowth}$ nor the initial population P_0 is known. However, $t_{regrowth}$ can be derived as follows.

The Probability Assay demonstrated that no regrowth would occur before $t = 11$ (See Table 5). Therefore

$$t = t_{regrowth} + t_{dormancy}$$

$t = 11$ days is known, and the variable of interest is $t_{regrowth}$

While neither $t_{regrowth}$ or $t_{dormancy}$ is known, the range of time $t_{regrowth}$ can be narrowed from $0 < t \leq 11$ to $0 < t \leq 4$ by data in the Probability Assay and an inferred value. Regrowth was never observed in the Probability Assay before day 11, and data from the control untreated groups in the MBC assays showed it took 4 days for a control group to grow from $\leq 10^5$ cells/mL to 10^7 cells/mL.

$$1 \leq t \leq 4 ; \text{ where } t = t_{regrowth}$$

To calculate k , the initial population concentration of 2×10^5 cells/mL can be used, which assumes the population is at log phase, and 2×10^7 cells/mL can be used for a final population concentration.

$$k = (\ln(P/P_0))/t$$

$$k = (\ln(2 \times 10^7 \text{ cells/mL})/2 \times 10^5 \text{ cells/mL})/96 \text{ hours}$$

k can be estimated to be 0.055 (See Appendix A, Section 1.1).

These criteria can be used to describe both the probability, and the quantity, of a persister phenotype subpopulation. As an example, consider two subpopulations. If a population of *B. burgdorferi* at 2×10^7 cells/mL is treated with the MBC of doxycycline, then for a time period of less than 20 days, there is a 0% chance of observing regrowth. Sometime after day 20, within the growth period of ($t_{regrowth} = 4$ days) there is approximately a 50% chance that a subpopulation of size 2×10^3 cells/mL will regrow. If a population of *B. burgdorferi* at 1×10^8 cells/mL is treated with the MBC of doxycycline, then for a period of 11 days or less, there is a 30% chance or greater that within a period of ($t_{regrowth} = 4$ days), a subpopulation of size 2×10^3 cells/mL will regrow.

Discussion

The addition of any substance, which causes alterations in the individual bacterium or small subpopulations of bacteria, can have profound effects on the bacterial population as a whole. These subpopulations, although genetically identical to the population, will uniquely respond to different external stimuli, such as those brought on by the antibiotics. These different responses can confer an advantage to survival of the bacterial population as a whole, when the population is subjected to adverse conditions [75, 114, 131, 132]. Although numerous

different genes and pathways in a bacterium may be expressed in response to a stimulus, if a simplified biphasic model is used to examine specific criteria, such as a growth or death response to a given stimulus like an antibiotic, then predictions about the probability of changes in state can be made [73, 133].

Since the subpopulations are genetically identical, and have no specific mechanism to inactivate antibiotics, this is fundamentally different from resistance. The bacterial subpopulations which correctly induce the genes that activate dormancy, while the larger population may get killed by an antibiotic, are termed persisters, a term coined by Dr. Joseph Bigger [113].

As the data suggests, *B. burgdorferi* populations can form phenotypically different subpopulations from antibiotic stress. The formation of these subpopulations is likely due to stochastic phenomenon, which can drive subpopulation transitions into to a stress-response state. This does not always occur, but as was seen in Table 5a-b, the higher the population density it is, the greater the chance that such a transition will occur. The results in Table 5c further suggest that simply lowering the cell count by diluting it may artificially increase the chance of persister development, and the phase of the population has an effect on persister development.

It has been shown with other bacteria [134-136] that oxidative stress or nutrient deprivation can induce a persister state. It is possible that a lack of readily available nutrients for some *B. burgdorferi* subpopulations, similar to what would occur at stationary phase, can induce persister development and

dormancy, and enable survival after the introduction of an antibiotic like doxycycline. Follow-up studies that will be performed by the Embers Lab using RNASeq will attempt to clarify the specific genes induced among *B. burgdorferi* survivors of antibiotic treatment.

APPENDICES

APPENDIX A: MATHEMATICAL MODELING

1.1 Mathematical Model to Quantify Persisters

The mathematical model was created to continue the Probability Assay, and to quantify the number in a subpopulation that transitioned from a growth state (or phenotype), to a dormant state, and to a growth state. The presence of persister subpopulations in a larger population can be described by:

For a given bacterial population p ,

$$p_f = (p_0)e^{kt}$$

where:

p_f = final population concentration

p_0 = initial population concentration,

k = constant

t = time

p is also composed of subpopulations p_1, p_2, \dots, p_n such that

$$p = \sum p_n$$

In addition, a given subpopulation p_n may change from one state to another state, where one state is a "survivor" that survives antibiotics and doesn't grow, while another state is a "non-survivor" state, that does not survive antibiotics but grows rapidly in normal media [131, 137]. During antibiotic treatment, bacterial subpopulations will either survive antibiotic treatment, or not, so the bacterial subpopulations can be described as

$$p_n = p_1 + p_2$$

where:

p_1 = subpopulation which transition to survivor state

p_2 = subpopulation which do not transition to survivor state

During treatment, exponential growth would not continue. However, a surviving subpopulation of bacteria of a larger population, which did not survive antibiotic treatment and experienced exponential growth after the removal of antibiotics could be modeled using exponential growth.

$$P_{\text{post-treatment}} = P_{\text{pre-treatment}} e^{kt}$$

$$P_{\text{regrowth}} = P_{\text{post-treatment}} e^{kt}$$

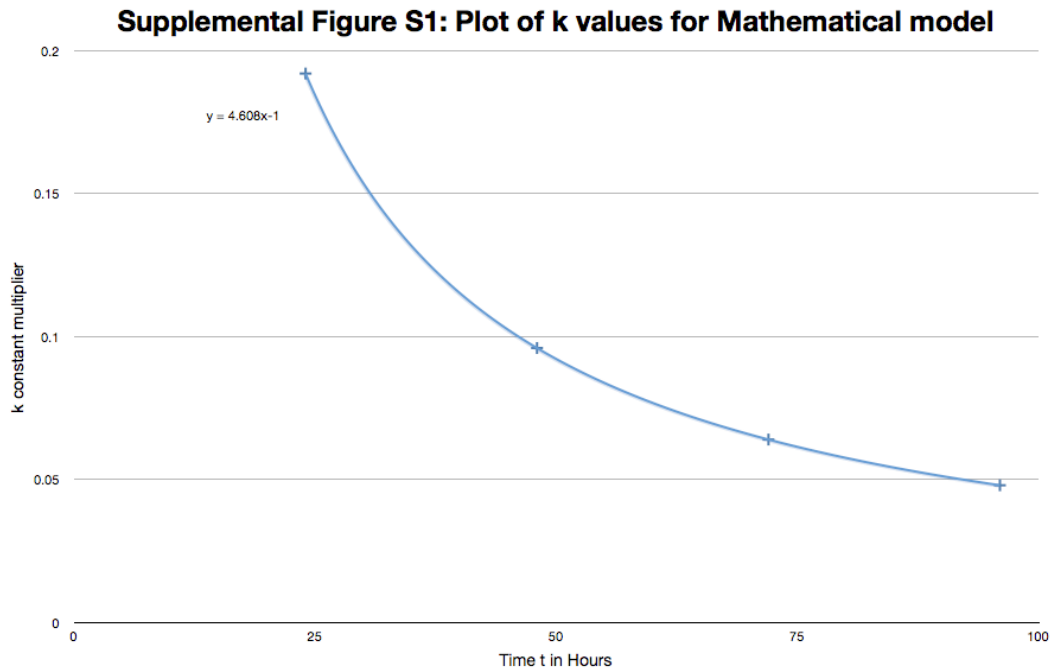
Where:

$$P_{\text{pre-treatment}} = p_1 + p_2$$

$$P_{\text{post-treatment}} = p_1$$

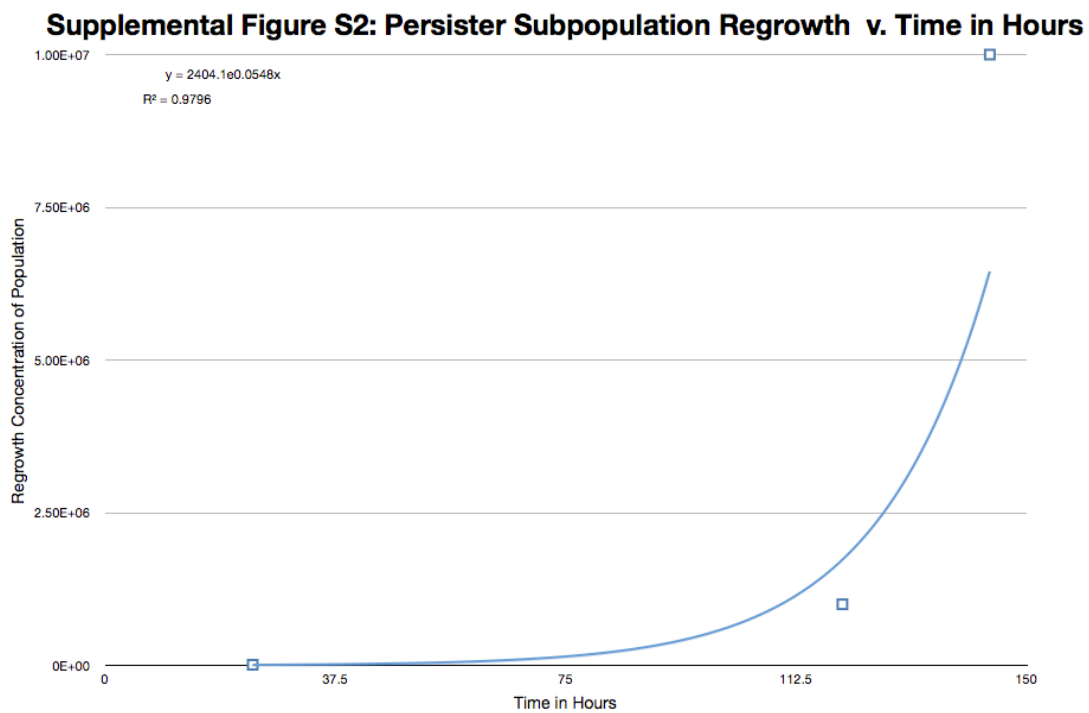
By examining the Pulse Dose Assay and the Probability Assay, the experimentally determined probability for regrowth after treatment during an assay of time t , $P(x) = \Sigma (\text{live})/(\text{dead})$ can be determined, and the amount of persisters in a given population can be quantified. Determining $P_{\text{post-treatment}}$ from $P_{\text{pre-treatment}}$ would be extremely difficult using the currently described methods, however, it can be calculated using P_{regrowth} , since the surviving persister regrowth population would have been dormant, but still alive, during antibiotic therapy.

As mentioned in Chapter 2, the k constant can be derived from the exponential growth formula. If the values for a range of $1 \leq t \leq 4$; where $t = t_{\text{regrowth}}$ are used and plotted, interestingly they perfectly match an exponential model for $y = 4.608/x$ (see Figure S1),



where $y = k$, and $x =$ number of hours

To further refine this approach, if the first cycle of the Pulse Dose Assay is fitted to a model, it shows $y = 2404e^{0.0548x}$, with $R^2 = 0.97$ (See Figure S2).



where the initial quantity of persisters = 2404

y = final population (regrowth of persisters)

x = time in hours

Therefore, a good approximation of k is $k \simeq 0.055$

1.2 Description of Differential Gene Expression Analysis

In a given sample, Partek® quantifies differential gene expression by modeling the number of reads that align to a transcript as a binomial distribution [138]. The model assumes that the transcript will be expressed the same in all samples. Significant expression is calculated by dividing the sum of the number of expressed reads, by the sum of total number of reads. The ratio is used to calculate the log-likelihood ratio, which is evaluated with the chi-squared test statistic. In the absence of replicates, Partek® still assumes that the reads are drawn from a binomial distribution, and proceeds accordingly [138].

n_s = number of reads that align to a transcript

N_s = total number of reads

For a given distribution, the null hypothesis $H_0 = \sum n_s / \sum N_s$, and the alternative hypothesis is defined as:

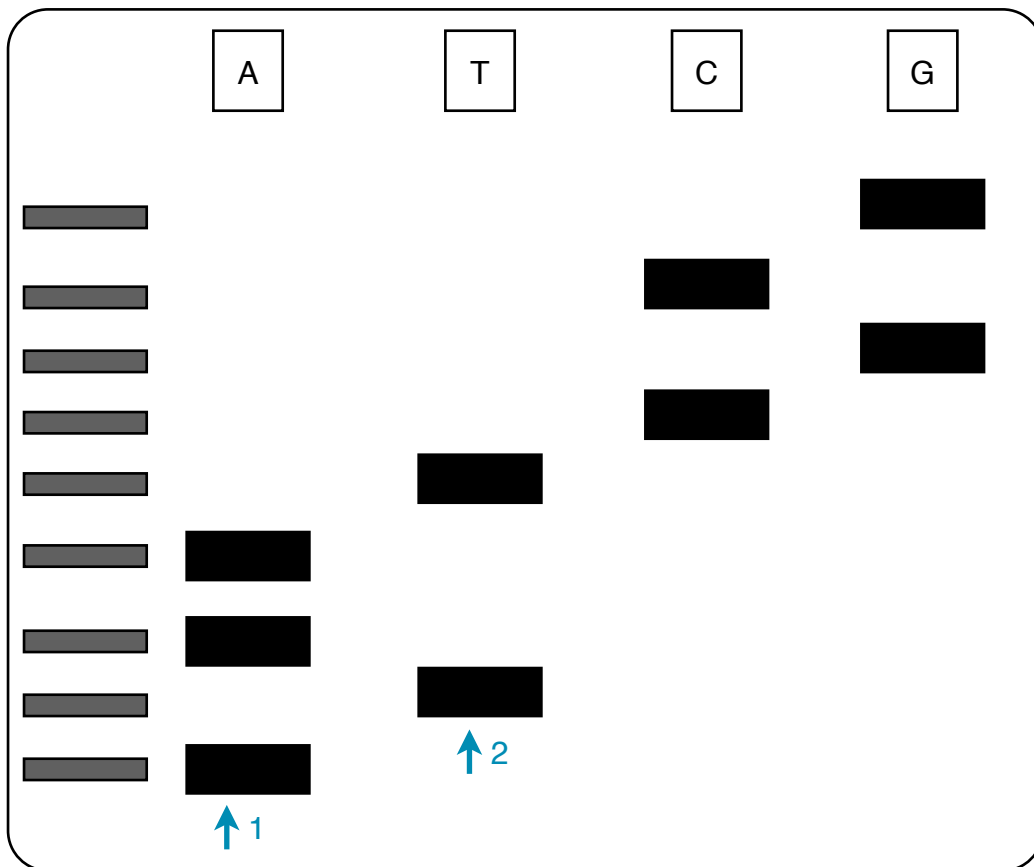
LogLikelihood = n_s / N_s , and $\sum n_s \log(\text{probability}(H_0) / \text{probability}(H_A))$

APPENDIX B: SEQUENCING THEORY AND METHODS

2.1 Sequencing Background

Research into determining the DNA sequence of an organism was advanced by Dr. Fred Sanger [139]. This sequencing method, commonly called “Sanger Sequencing” or “capillary sequencing”, can best be explained by assuming the genome to be sequenced is a black box, where the precise sequence is unknown, but continuous. Deoxy-nucleotide triphosphates, along with di-deoxynucleotide triphosphates are added to a primer with a polymerase to the complementary strand of a sequence of DNA of interest. The polymerase will randomly incorporate either a deoxy-nucleotide or a di-deoxynucleotide, resulting in either elongation of the DNA or termination, since the di-deoxynucleotide lacks a hydroxyl to continue replication. These stretches of DNA can then be visualized on an agarose gel with a visibility marker of some kind to determine the sequence of DNA. The shorter stretches are at the beginning of the sequence, and the longer stretches of DNA are at the end of the sequence, and the sequence can then be determined by reading up a gel [139]. For example, a stretch of DNA sequenced by Sanger Sequencing would read on a Gel similar to

Supplemental Figure S3.



Supplemental Figure S3. Illustration of Sanger Sequencing. An unknown stretch of DNA is amplified with PCR and a di-deoxynucleotide triphosphate terminator molecules. Polymerases that incorporate a ddATP will halt the DNA stretch elongation, yielding the stretch by arrow 1. A polymerase that incorporates a dATP, then a ddTTP will halt elongation, yielding the stretch by arrow 2, and so on. If the DNA from the PCR reaction is then run on a gel, the previously unknown sequence will be: ATAATCGC

2.2 Next Generation Sequencing and RNASeq

These advances were succeeded by several different techniques to accomplish sequencing, called Second Generation Sequencing (SGS), and then Next Generation Sequencing (NGS). Interest in RNASeq, or the accurate and

quantitative analysis of an organism's total mRNA transcripts (organismal transcriptome) began to increase as well [93]. An RT-PCR could show expression for a single gene, or potentially a small set of genes, and the microarray was used to broaden the number of potential genes that could be simultaneously studied. A microarray shows relative abundance of expression either in DNA, or cDNA from mRNA, by comparing hybridization of a sample to a probe, and measuring the corresponding intensity over large gene sets from an organism, including the organism's entire transcriptome (if using cDNA) [140].

Several challenges remained if a microarray was used for RNASeq analysis. Typically, for both RT-PCR and microarrays, the sequence of the gene of interest had to first be known. Similar probe sequences could cross-hybridize and confound results, or limit the analysis. Microarray results were recorded by differing intensity output, which could be subject to error or false positives [98, 141]. With newer NGS methods, it became possible to conduct highly accurate and quantitative RNASeq transcriptome experiments.

Currently, there are several companies which can perform NGS sequencing, such as Illumina/Solexa, Helico, Life Technologies, Roche/454, and Pacific Biosciences. While the overall methodology and theory is similar for all platforms, differences arise in how the sequencing is specifically carried out. During experiments for this dissertation, the Life Technologies Ion Torrent Personal Genome Machine (PGM) NGS platform was used for RNASeq analysis, but a comprehensive view of the technology and theory is important to

understand RNASeq and NGS. Therefore, what follows will be an outline of the methods and theory in NGS and RNASeq, an extensive explanation of Life Technologies' Ion Torrent PGM platform, and brief synopses other sequencing platforms. Sample preparation will be only mentioned in this section, but Chapter 2 will cover a both an overview of sample preparation for sequencing on any platform, and an extensive description of the sample preparation used with Ion Torrent PGM.

2.3 NGS Platforms

At its most basic level, NGS encompasses the following steps: isolation of genetic material, attachment of material to a surface for analysis, determination of the genetic sequence, and output with error correction and quality assessment. To start, either DNA or total RNA must be isolated from the source to be studied. Following isolation and preparation of the sample, one of two methods is used to create the template which will be used for analysis: clonal amplification of the template, or the use of a single molecule template [142]. Illumina/Solexa, Roche/454, and Life Technologies use clonal amplification, with variations on how it is implemented. In Illumina kits, a process known as solid-phase amplification occurs when the DNA library is fragmented randomly, ligated to adapters with or without PCR (depending on the Illumina kit used), and then attached to primers, which have been fixed to a glass slide [142-146]. Clusters of template are created with bridge-amplification by hybridizing the fragmented

sample to primers which are attached to a glass surface, and then cyclically elongating the primer with a hybridized template, detaching the template during a denature step, and either hybridizing to an attached elongated primer, or a template [142, 144, 145]. Life Technologies and 454/Roche use a different process to create clonally amplified templates called emulsion PCR, or emPCR (See Supplemental Figure 2B) [142, 147]. During emulsion PCR, beads, templates, and primers for the adaptors at the ends of the templates are captured into an emulsion bubble (for example, oil and water), with a low template to bead ratio to minimize the chance more than one template will be in any one emulsion [141]. Next, the template/primer strands attach to sites on the beads with complementary primers attached, and PCR amplification extends the template copy on the bead [141, 142]. Then the emulsion stops, and this cycle continues until some of the beads have numerous copies of the template attached. At this point, the beads are selected and NGS proceeds with the template (PGM), the beads are deposited into wells before proceeding with the template (454/Roche), or the beads can be attached to a glass slide (others) [141, 142, 148, 149]. Alternatively, Helicos and Pacific Biosciences use a single molecule template method, whereby the sample library is randomly fragmented, then attached to a templates or primers group. The primer group has been previously attached to a glass surface (Helicos), or attached to a primer, then added to a polymerase, which has been previously attached to a surface (Pacific Biosciences) [143, 150].

Following sample preparation and template preparation, determination of

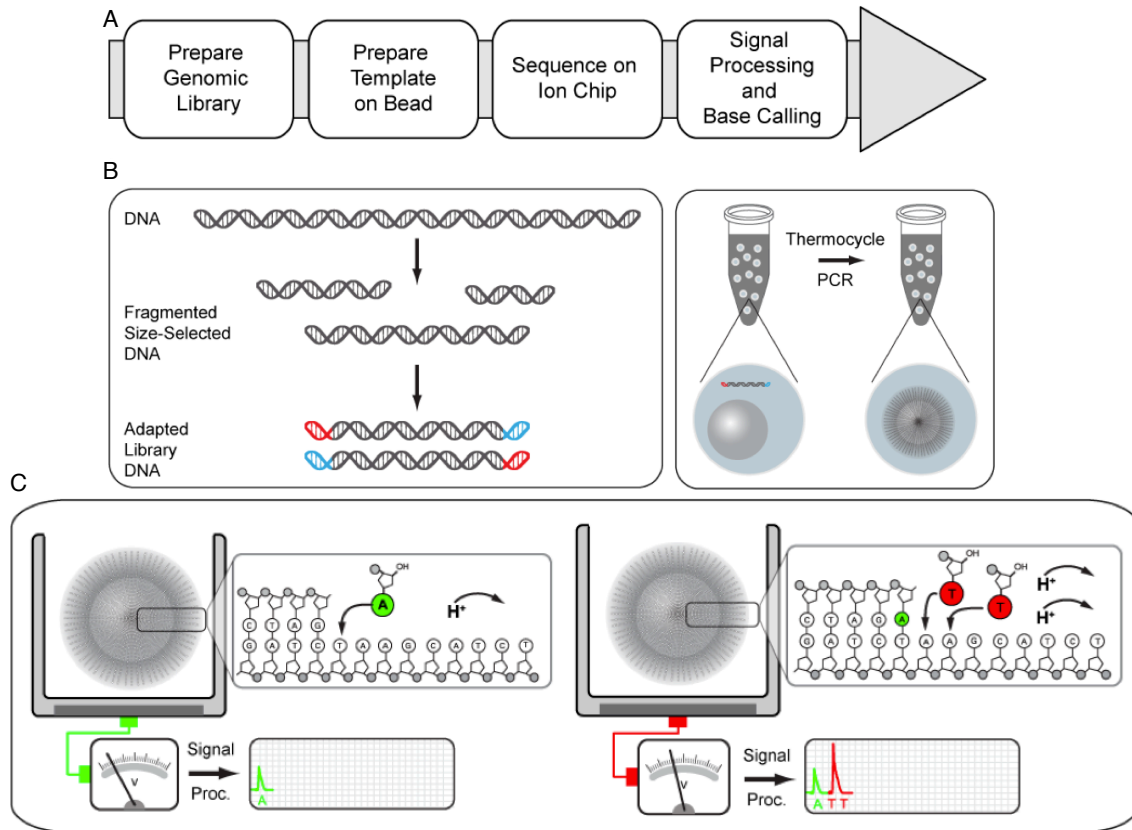
the template sequence takes place, and is broadly categorized into cyclic reversible termination, single nucleotide sequencing, and semiconductor sequencing. One sequencing method in use by Illumina/Solexa builds on the technology initially used in Sanger sequencing, and uses cyclic reversible terminator molecules. Essentially, replication proceeds with incorporation of the nucleotide by a polymerase and cleavage of the molecule for incorporation, the release of a fluorophore molecule (specific wavelength of light to indicate A,T,C,G, usually red, green, blue, or yellow) and temporary termination of elongation. This step is followed by wash and cleavage steps to remove fluorophores, nucleotides, and terminating groups so elongation can proceed [142]. Cycles of these steps are continued until the sequence of the clonal templates has been recorded. An alternative method used with cyclic reversible terminators, is in use by Helicos with a single molecule template. Similar to Illumina/Solexa, the nucleotides with a linked fluorophore are singly incorporated into the template. A bulky group attached to the fluorophore serves as an inhibitor to the polymerase during instances of repeats in the genome [142, 151]. Subsequent wash and cleavage steps remove the dye and blocking group, and the cycle continues until the sequence of the template has been recorded. Roche/454 sequencers have a clonally amplified template, but utilize single molecule sequencing. A bead, attached with numerous copies of the clonally amplified template, has a polymerase sequentially elongate the template, and after each reaction the inorganic phosphate reacts with ATP sulfurylase and

luciferase to create a specific wavelength of light, enabling the sequencing of individual nucleotides [152]. Despite many advances in sequencing, visualizing the polymerase elongate a stretch of DNA, or “real-time” sequencing, is not an easy task. The background fluorescence of existing fluorescently-linked nucleotides can interfere with readings (other technologies, even pyrosequencing, have a wash step). Creating a small enough space that can hold and capture readings from bases as they were incorporated (base calling) accurately can be difficult and prohibitively expensive [150]. Pacific Biosciences was able to overcome both of these obstacles by anchoring the polymerase to a chamber floor, and designing a filter to selectively filter out all light except light emitted from a cleaved fluorescent marker during incorporation [153, 154].

2.4 Ion Torrent

Semiconductor sequencing is a novel sequencing method, used by Life technologies, which eliminates the need for light to determine a DNA sequence. Chips are prepared with individual wells for libraries and a metal-oxide plate to sense the addition of nucleotides by changes in pH, called Complementary Metal-Oxide Sensing (CMOS) [149]. As mentioned previously, an emPCR clonally amplified template on a bead is deposited into each well (Supplemental Figure S4).

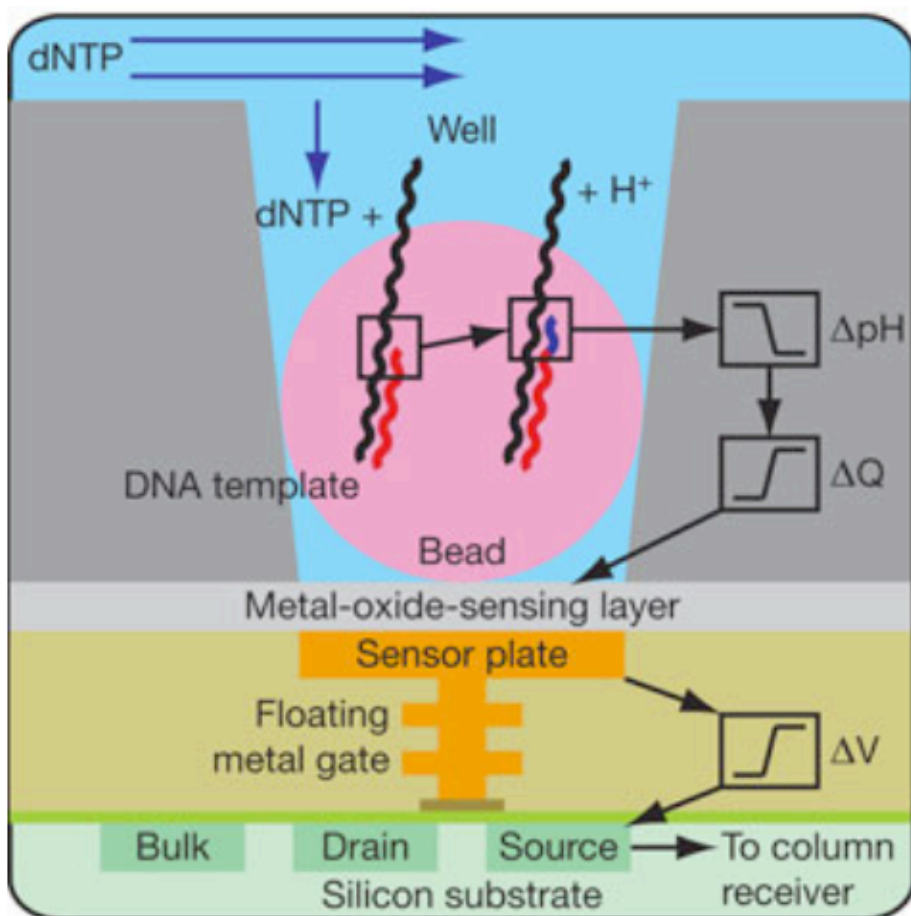
Supplemental Figure S4. Ion Torrent Processing and Sequencing Overview



Supplemental Figure S4. Ion Torrent Sample Processing and Sequencing Overview. (a) Overview of Ion Torrent sequencing Pipeline (b) cDNA or DNA is randomly fragmented, then attached to adapter linkers, which are then attached to beads for emPCR (c) Sequencing is detected by changes in pH as nucleotides are sequentially added to the cDNA libraries, which are attached to emPCR beads. [149] Used with Permission.

During sequencing, nucleotides are added in a set sequence, similar to pyrosequencing, and the sequence is determined by corresponding changes in pH (See Supplemental Figure S4, S5). To accommodate for repeats, corresponding increases in pH are detected, and software algorithms correct for phase correction and base calling [149].

Supplemental Figure S5. Overview of Ion Torrent Complementary Metal-Oxide Sensing (CMOS) Sequencing



Supplemental Figure S5. Schematic overview of the sequencing process. Nucleotides are added to beads inside wells, and changes in pH are detected by the metal-oxide sensor, which relays information to the sequencing equipment. [149] Used with Permission.

During the sequencing process, analytical software corrects for phase variation, or differences in sequencing across multiple clonally amplified libraries, and normalizes signal intensity [149]. The software also excludes low quality reads, and assigns each read a quality score, which is based on the Phred algorithm [149, 155]. The Phred algorithm assigns a quality score to each

base in each read, and is used both in overall assessment and in analytical assessment of the reads to determine how precise the sequencing was performed [156].

2.4.1 RNASeq on Ion Torrent

RNASeq can also be performed on these platforms by converting transcriptome libraries to cDNA. Ideally, RNASeq should be performed when the strand orientation can be preserved, since RNA can be copied in either the forward or reverse direction. Recent developments in NGS have allowed this to happen, and several platforms such as Ion Torrent and Illumina HiSeq are well suited to this task.

When performing RNASeq using NGS platforms like Ion Torrent, initial preparation differs slightly, and most methods center on processing the RNA, then creating a cDNA library from the RNA. Interestingly, a new study has proposed directly sequencing RNA as it is converted to cDNA [157].

With RNASeq on Life Technologies Ion Torrent PGM, total RNA is first extracted from a sample, and then ribosomal RNA is depleted from the total RNA. Since transcript RNA is the desired source of information, it is usually recommended to deplete ribosomal RNA (rRNA) from the total RNA sample, but this can be skipped if the total RNA content is high enough [158]. After rRNA depletion, the rRNA-depleted RNA (abbreviated here rdRNA) is then randomly fragmented to create a nonbiased template of the rdRNA [142, 148, 149]. The

rdRNA is then size selected for an approximately binomial distribution centered at 150 base pairs [159-161]. Size selection occurs both at the RNA and cDNA steps. Briefly, it was discovered [159] that if magnetic beads were coated with a molecule containing an exterior-facing carboxyl group, a nucleic acid, such as DNA, would bind to the magnetic bead. If the bead-DNA complex were washed with a polyethylene glycol and salt solution, and then water, the DNA would dissociate [159]. Furthermore, if a water-and-ethanol solution were used instead of pure water, a specific base pair size range of nucleic acids could be isolated, depending on the concentration of the ethanol [160].

Next, the purified and size-selected RNA fragments are ligated to adaptor molecules, and then converted to cDNA with a reverse transcriptase (RT) reaction [161, 162]. The cDNA is selected for a specific size range following the same process as before. Finally, the cDNA is amplified with a high-fidelity PCR reaction, and size-selected, before proceeding with NGS as mentioned previously [161]. As an example, Supplemental Figure S6 and the following section show the preprocessing pipeline used with samples for experiments for this Dissertation.

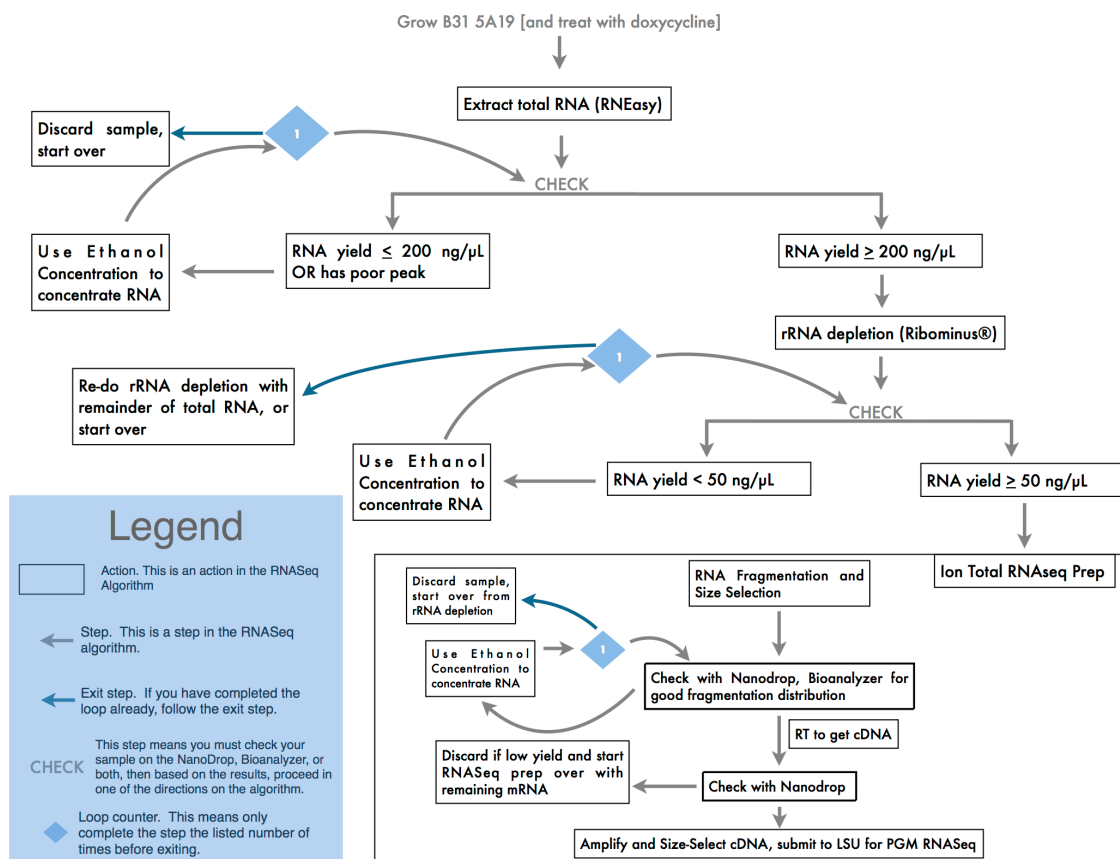


Figure S6. RNASeq Work Flow for Sample Preparation

2.4.2 RNASeq Sample Preparation

B. burgdorferi cultures were grown to 5×10^7 cells/mL for the control, treated, and treated/regrown groups in 50 mL conicals in BSK-II media. For the control group, total RNA was extracted, while the treated and treated/regrown groups were treated with $50 \mu\text{g/mL}$ doxycycline for 5 days. On day 5, the treated group total RNA was extracted, while the treated/regrown group was resuspended in doxycycline-free media, and monitored for regrowth. When the culture regrew to the initial concentration, total RNA was extracted. In all cases, RNA was kept in a -80° freezer when not being processed, and freeze-thaw

cycles were kept to a minimum. After total RNA extraction, the total RNA was assessed for concentration and purity with a NanoDrop™ 2000 spectrophotometer (ThermoFisher Scientific). A concentration greater than 300 ng/μL should be obtained before proceeding beyond this step. Next, for all groups, ribosomal RNA was depleted using Invitrogen Ribominus kits. These kits function by binding to the 16S and 23S RNA with a streptavidin-biotin linker, which is then bound to a magnetic bead. The linkers in turn have locked-nucleic-acid monomers, which raise the bound molecule's melting temperature, and allow removal of ribosomal RNA [163]. The rRNA-depleted RNA for all samples should have a concentration of at least 50 ng/μL before proceeding. The RNA for all groups was then processed according to the Life Technologies Ion Total RNA-Seq kit v2 instructions [161]. Briefly, the RNA was converted into a cDNA library, selected for a size range, which matched a binomial distribution, and then sequenced with the Ion Torrent PGM.

2.5 Alignment Algorithms

Several alignment algorithms exist for bioinformatics. The purpose in this Appendix is not to provide an exhaustive description, but instead to provide a few relevant descriptions to illustrate how sequences are aligned to a reference genome. As mentioned previously, the purpose in alignment algorithms is to provide an alternative to a “brute force” approach of trying every possible alignment at every possible position.

2.5.1 Algorithm Theory Overview

At its most basic level, an algorithm is a series of defined steps that takes an input value and provides an output value [164]. These days, algorithms are usually associated with computers, but the analysis and problems that algorithms are used to solve is applicable across many disciplines. In Biology and Bioinformatics, algorithms are a useful tool that can be used to analyze vast datasets, such as organismal genomes. Since a small genome is thousands of bases long, it becomes difficult to impossible for someone to simply look to find patterns of gene expression, Single Nucleotide Polymorphisms (SNP's), etc. An algorithm can therefore be written to undertake the analysis, and initially it can be as simple as:

Input data -> find match for a gene in genome -> return location of match

Subsequent steps refine the algorithm, and subdivide the tasks that will be accomplished.

A large portion of writing and analyzing algorithms is a component called the cost, which includes both the computational cost, or how many resources are required to complete the task, and the running time, or how many seconds, hours, or days, it will take to complete the task [164]. For simplicity, they will both be merged together into the running time cost. The cost of an algorithm can be

broken down into how long the each step in a series of steps takes, and is reported using a convention called “Big-Theta” relationship [164, 165].

Given a series n steps for the function, the series of steps can be defined as $f(n)$. How quickly $f(n)$ grows determines the cost in the function, and is bounded by in the upper range by a function $O(f(n))$, and in the lower range by $\Omega(f(n))$. The faster a function grows, the higher the cost and the longer the running time will be. To put it another way, between an upper-bound “worst case” growth, $O(f(n))$, and a lower-bound growth and $\Omega(f(n))$, there exists a function $\Theta(f(n))$ which describes the asymptotic theta relationship of the growth of algorithms [164, 165]. Generally speaking, it is desirable to have the growth either logarithmic, or possibly linear, and not exponential, unless the input size for the algorithm is very small. An algorithm that is described by a function with a growth of $50 \cdot \log n$ is better to have than $2x^2$, unless the input is very small, since exponential growth would mean an exponential asymptotic increase in the cost and running time.

A concrete example can be visualized with the dataset that was used for Chapter 3, Table 4. The data from the RNAseq experiment had previously been split into two datasets. The genes were from the same organism, but one dataset had genes that were 2-fold upregulated in the control versus the treated sample, and the second had genes that were significantly expressed in the control versus the treated sample. The task was the genes, which were 2-fold upregulated, needed to be matched with the genes which were significantly

expressed between datasets, but there were over 1500 genes. One approach was to simply write a short python program to tell the computer to start at the first gene in the 2-fold upregulated set, check every gene in the significantly expressed set, and report if there is a match. The pseudocode, with an approximated analysis, is in Table ST1.

Table ST1. Pseudocode and Cost for Match Algorithm		
Task	Value	Result
1. for each gene in the first set:	let: FirstTask = a	
2. for each gene in the second set	let: SubTask = b , with size n	
check all genes against the first gene, if match, return "Match"		For task a , and subtask b of task a , it takes $T(n^2b)$ time to finish a tasks
		$T(n) = \Theta(n^2)$

Using intuition and the Master Theorem [164] to analyze the algorithm, it is a "brute force" approach with Θn^2 running time. Thankfully the computer was able to match all of the genes without crashing. As an alternative, consider if it was noticed before starting that there were roughly twice as many genes in the second group as the first group. Therefore, while the numbering for the genes in either group was not consistent, if there was a match for the genes in both groups, it would appear about every other gene. This can be visualized as instead of taking one gene, and checking every other possible gene in the second group, taking one gene, and checking a narrow window of genes in the second group, because it is not possible for the other genes outside of that window to match. To ensure the accuracy of the program, the "window" could be expanded to 5, or even 6 genes, and is displayed in Table ST2.

Table ST2. Pseudocode and Cost for Revised Match Algorithm		
Task	Value	Result
1. Sort each gene in the first and second set into a numbered array.	$D(n)$	$D(n)$ time to divide the problem into subproblems
2. for each gene in the first set	let: FirstTask = a	
check genes in a window of size x to $x+5$ for match, if match, return "Match"	each subTask size = $1/b$, $b \ll n$	
return match	$C(n)$	$C(n)$ time to return result
		Time $T(n) = 2T(n/5) + D(n) + C(n) = \Theta(\lg n)$

This algorithm is much more efficient, and runs in $\log n$ time. For larger datasets, this difference is critical, and could mean the difference between a computer being able to, or not being able to, process a task.

2.5.2 Burrows Wheeler Transform

In 1994 Burrows and Wheeler, working at HP, published a paper on a new compression method [166]. They noticed that in stretches of text runs of similar characters frequently appear together if the characters were lexicographically sorted. Building on this observation, if the frequency of these runs of characters could somehow be stored, then the original text could be retrieved in its original form [166]. Their method utilized a suffix tree, or an ordered set of partial strings of the original string, which could be traversed to recreate the original string. Subsequently, it was observed that the position in the string could be noted as well, and the suffix tree could be adjusted to accommodate for mismatches to locate and align strings to a reference string [167]. This method, called the

Burrows Wheeler Transform, is used in several Bioinformatics software platforms.

2.5.3 Seed Matching

A second method for matching a string to a reference genome utilizes the idea that if a string matches a reference genome, there may be mismatches between both, but if the string and genome are divided into smaller and smaller pieces, there will be a segment that matches exactly [168]. Once a match is found, it can then be mapped back to the reference genome using techniques such as a Smith-Waterman algorithm [168, 169]. A common and well-known application of this algorithm is the online tool National Center for Biotechnology Information Tool, BLAST.

APPENDIX C: RNASeq Database BTech

3.1 Overview

The *Borrelia* RNA Sequencing Technology Database, or BTech, serves two main purposes. First, it acts as a repository of the *B. burgdorferi* genome that is easily searchable and accessible. Second, it provides useful gene annotation data for *B. burgdorferi* under untreated, treated, and treated/regrown doxycycline conditions. This data was obtained from the RNASeq experiments that have been discussed previously.

To create the database, gene information for the *B. burgdorferi* B31 genome, or RNASeq expression data for each gene, was exported as a comma-separated-value (.csv) spreadsheet. The spreadsheet was then uploaded to a Drupal (version 7.x) installation at the website <http://borrelia.rna.technology> using the required Drupal module. The data was displayed either using core Drupal modules, or add-on modules as noted. The data was sorted for display with a customized Drupal module, which used php scripts to send MySQL queries. A limited interface of the genomic data can be viewed by anyone navigating to the website, however, a username/password login is required to fully search and

access the *B. burgdorferi* genomic data on the website. At this point in time, the analysis is preliminary due to a lack of replicated data points, however, the analysis can still be used to make predictions to determine which genes are up- and down- regulated, and the framework that has been created can be easily reused with future analyses.

3.2 Drupal Overview

For an in-depth description of Drupal, it is highly recommended to go to <https://www.drupal.org/> or browse numerous available books on the topic. This section will simply provide a broad overview of the organization and framework that Drupal uses, for the purpose of detailing how BTech was set up and functions. Drupal is a content management system (CMS) that allows the rollout, distribution, and management of content on the Internet. It uses a core installation, with installable modules that enhance the functionality and capabilities of the website.

Drupal manages content on a website with nodes of data. Properties to change the behavior and appearance of nodes can be applied to individual nodes, or all nodes across the website. When a user accesses a webpage, they are accessing a node of data that displays according to the properties that have been assigned to it.

3.3 Adding Content to Drupal

To begin, a hosted website was created at <http://rna.technology> with a “vanilla” Drupal version 7.x installation. Drupal 7 was chosen due to greater module availability, and timing; when the database was created, Drupal version 8 was nearly set to launch. Due to time constraints, a customized theme was not used, but a modified version of Barto (renamed “Bartik”) was applied to the website instead.

There are several templates of webpages that can be used for the upload process. Essentially, all that needs to be done is to create a webpage based on the way that the gene will be displayed (See Supplemental Figure S6), with fields for data to be added like description, gene expression, p value, etc.

BB_genome

Home » Administration » Structure » Content types » BB_genome

Default Teaser Bb Data ListView

Content items can be displayed using different view modes: Teaser, Full content, Print, RSS, etc. Teaser is a short format that is typically used in lists of multiple content items. Full content is typically used when the content is displayed on its own page.

Here, you can define which fields are shown and hidden when content is displayed in each view mode, and define how the fields are displayed in each view mode.

FIELD	LABEL	FORMAT
Body	<Hidden>	Default
plasmid name	Above	Default
Description	Above	Default
CoDSeq	Above	Default
Gene Number	Above	Default
B31P v B31 2 Fold	Above	Default
B31PG v B31 2 Fold	Above	Default
Protein ID	Above	Default
BbTag	Above	Link
p value	Above	Default

1 234.12345679
Display with prefix and suffix

Supplemental Figure S7. Fields in Drupal. After creating a page from a template, or creating a new template, fields, that contain data for each node, are added with text, number, or other values, and how they are displayed can be customized.

Whether data was gene information from the National Center for Biotechnology Information (NCBI), RNASeq experiment gene expression data, or taxonomy data, unless information was manually typed into the website, any data had to be converted to a .csv file with spreadsheet software. Microsoft® Office usually worked the best for unknown reasons, although OpenOffice and Apple® Numbers also worked reasonably well. To upload the data, the Feeds module (see Supplemental Figure S7 A), which was designed for RSS feeds, but was found to be the most usable and stable for .csv file import, was installed and configured to import gene values, taxonomy values, and RNASeq experiment expression values. The Feeds module created a feeds importer (see Supplemental Figure S7 B), which had several configuration options. The best setup for the importer was found to be to upload the .csv file to a directory on the webserver, which Drupal would consider “local”. Turn off periodic import, and set the parser to “comma”. The setting “import on submission” caused unpredictable outcomes if turned off, and should be left on, unless a dire need dictated otherwise. The processor for the feeds importer should generally be “Node”; if set to “User” this would cause changes to the user file, and if set to “Taxonomy”, this would cause changes to the taxonomy terms. Mapping (see Supplemental Figure S7 C) assigns which fields of the node will receive the data after the .csv file is successfully imported.

A

FEEDS

ENABLED	NAME	VERSION	DESCRIPTION	OPERATIONS
<input checked="" type="checkbox"/>	Feeds	7.x-2.0-alpha8	Aggregates RSS/Atom/RDF feeds, imports CSV files and more. Requires: Chaos tools (enabled), Job Scheduler (enabled) Required by: Feeds Import (enabled), Feeds News (disabled), Feeds Admin UI (enabled)	Help Permissions
<input checked="" type="checkbox"/>	Feeds Admin UI	7.x-2.0-alpha8	Administrative UI for Feeds module. Requires: Feeds (enabled), Chaos tools (enabled), Job Scheduler (enabled)	Configure
<input checked="" type="checkbox"/>	Feeds Import	7.x-2.0-alpha8	An example of a node importer and a user importer. Requires: Feeds (enabled), Chaos tools (enabled), Job Scheduler (enabled)	
<input type="checkbox"/>	Feeds News	7.x-2.0-alpha8	A news aggregator built with feeds, creates nodes from imported feed items. With OPML import. Requires: Features (missing), Feeds (enabled), Chaos tools (enabled), Job Scheduler (enabled), Views (enabled)	

B

Home » Administration » Structure

Create one or more Feed importers for pulling content into Drupal. You can use these importers from the [Import](#) page or – if you attach them to a content type – simply by creating a node from that content type.

[+ Add importer](#) [+ Import importer](#)

NAME	DESCRIPTION	ATTACHED TO	STATUS	OPERATIONS	ENABLED
Add data for taxonomy	Trt vs ctrl datapoints	[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
add more data for taxonomy	regn vs ctrl	[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
gene DB update	update for incorrect genes	[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
Import Gene Data	Import Gene Data to Website	[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
update gene data	add RNASeq data	[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
update gene pg data		[none]	Normal	Edit Export Clone Delete	<input checked="" type="checkbox"/>
Node import	Import nodes from CSV file.	[none]	Default	Override Export Clone	<input checked="" type="checkbox"/>
User import	Import users from CSV file.	[none]	Default	Override Export Clone	<input checked="" type="checkbox"/>

[Save](#)

C

Basic settings

Attached to: [none] [Settings](#)

Periodic import: off

Import on submission

Fetcher [Change](#)

File upload [Settings](#)

Upload content from a local file.

Parser [Change](#)

CSV parser [Settings](#)

Parse data in Comma Separated Value format.

Processor [Change](#)

Node processor [Settings](#) [Mapping](#)

Create and update nodes.

Mapping for Node processor [Help](#)

Define which elements of a single item of a feed (= Sources) map to which content pieces in Drupal (= Targets). Make sure that at least one definition has a *Unique target*. A unique target means that a value for a target can only occur once. E. g. only one item with the URL <http://example.com/content/1> can exist.

[Show row weights](#)

SOURCE	TARGET	TARGET CONFIGURATION	
<input type="checkbox"/> geneName	Title	Used as unique.	<input type="checkbox"/> Remove
<input type="checkbox"/> geneNumber	Gene Number		<input type="checkbox"/> Remove
<input type="checkbox"/> plasmidName	plasmid name		<input type="checkbox"/> Remove
<input type="checkbox"/> CDS	CoDSeq		<input type="checkbox"/> Remove
<input type="checkbox"/> geneDescription	Description		<input type="checkbox"/> Remove
<input type="checkbox"/> proteinName	Protein ID		<input type="checkbox"/> Remove
<input type="checkbox"/> plasmidID	plasmid	Search taxonomy terms by: Term name	<input type="checkbox"/> Remove

- Select a target -

The name of source field. The field that stores the data.

Supplemental Figure S8. Overview of Import Process. (A) After a target webpage has been created, the Feeds module will can assign data from a .csv file to the page and create new pages if needed. (B) The Feeds module can create feeds importers for different tasks with different configuration options. (C) The mapping settings maps values from the .csv file to fields on the node. In this case, for each row in the .csv file, the value in the “geneName” column will map to the “Title” field, the value in the geneNumber column will map to the “Gene Number” field, etc., in a new Drupal node.

For example, “Gene Name” column in the .csv file will be assigned to the “Title” field on each new node that is created. It is **strongly urged** to backup the entire database and website before attempting any large .csv file import. After the import process, the nodes may need some minor modifications.

3.4 Indexing and Searching

Among other jobs, cron tasks create an index of field terms for each node. However, to reliably access a node by a specific search term, it is much better to create a specific taxonomy term for it, which assigns a taxonomy term within the search index to the field and the node. Taxonomy terms can be created as a group or individually, and can be added manually (see Supplemental Figure S8 A), or en masse using the Feeds module (see Supplemental Figure S8 B). As an example, the term “ospA” was assigned to the group “Bb transcription”, and once indexed, will display the node associated with it if ospA is typed in a search field (see Supplemental Figure S8 C).

A

Taxonomy My account Log out

Home » Administration » Structure

Taxonomy is for categorizing content. Terms are grouped into vocabularies. For example, a vocabulary called "Fruit" would contain the terms "Apple" and "Banana".

[+ Add vocabulary](#)

[Show row weights](#)

VOCABULARY NAME	OPERATIONS
+ B31P v B31	edit vocabulary list terms add terms
+ B31PG v B31	edit vocabulary list terms add terms
+ Bb GOI	edit vocabulary list terms add terms
+ Bb regulatory	edit vocabulary list terms add terms
+ Bb rpos	edit vocabulary list terms add terms
+ Bb sort	edit vocabulary list terms add terms
+ Bb transcription	edit vocabulary list terms add terms
+ Borrelia Genome	edit vocabulary list terms add terms
+ Borrelia opp	edit vocabulary list terms add terms
+ Borrelia osp	edit vocabulary list terms add terms

B

Add data for taxonomy EDIT EXPORT CLONE DELETE

Home » Administration » Structure » Feeds importers

Basic settings

Attached to: [none] [Settings](#)

Periodic import: off

Import on submission

Fetcher [Change](#)

File upload [Settings](#)

Upload content from a local file.

Parser [Change](#)

CSV parser [Settings](#)

Parse data in Comma Separated Value format.

Processor [Change](#)

Node processor [Settings](#) [Mapping](#)

Create and update nodes.

Select a processor [Help](#)

Node processor [Select](#)

Create and update nodes from parsed content.

Taxonomy term processor [Select](#)

Create taxonomy terms from parsed content.

User processor [Select](#)

Create users from parsed content.

[Save](#)

C

BTech My account Log out

Search Genome

Home About Login Genome

Home

BB_A15

[View](#) [Edit](#) [Manage display](#) [Revisions](#) [Log](#)

Submitted by admin@ on Sun, 04/13/2014 - 16:00

plasmid name:
lp54

Description:
outer surface protein A %28OspA%29

CoDSeq:
9457--10278

Gene Number:
gene12

B31P v B31 2 Fold:
B31P down vs B31

Supplemental Figure S9. Taxonomy Terms. (A) Taxonomy terms can be in groups, or individually, and can be manually entered into Drupal. (B) Alternatively, the Feeds module can bulk-import and assign taxonomy terms. (C) The taxonomy term "ospA" has been assigned to the node (and gene) BB_A15, and when "ospA" or "BB_A15" is typed into a search, the node (webpage) appears as a search result.

3.5 Displaying Data

Data is displayed either using features contained in the Drupal installation like views (see Supplemental Figure S9 A), which automatically create a list table, or as a Google® Chart [170] image. The Google® Chart API (see Supplemental Figure S9 B), which is installed as a Drupal module, allows nodes to either present values, or return values from MySQL queries via php scripts, in a chart image. The chart is then displayed on the node webpage in a human-readable form (see Supplemental Figure S9 C).

A

Bb Genome

Title	plasmid name	Gene Number	Description	CoDSeq	Protein ID
BB_0001	chromosome	gene0	hypothetical protein	105--677	NP_212135.1
BB_0002	chromosome	gene1	beta-N-acetylhexosaminidase	768--1796	NP_212136.1
BB_0003	chromosome	gene2	pseudogene	Null	null
BB_0004	chromosome	gene3	phosphoglucosyltransferase	3419--5188	NP_212138.2
BB_0005	chromosome	gene4	tryptophanyl-tRNA synthetase	5251--6306	NP_212139.2
BB_0006	chromosome	gene5	membrane protein	6309--7397	NP_212140.2
BB_0007	chromosome	gene6	hypothetical protein	7458--8315	NP_212141.1
BB_0008	chromosome	gene7	hypothetical protein	8427--9197	NP_212142.2
BB_0009	chromosome	gene8	hypothetical protein	9202--10206	NP_212143.1
BB_0010	chromosome	gene9	holo-acyl-carrier protein synthase	10203--10577	NP_212144.1
BB_0011	chromosome	gene10	hypothetical protein	10581--11420	NP_212145.1

B

Body

Below is a chart of the genome, and the distribution of linear plasmids, circular plasmids, and the linear chromosome. The major pathways as defined by Fraser et al (Nature 1997) are also listed. Genes and plasmid views can be displayed by searching in the search field in the upper right hand corner, or selecting the desired link from the list below.

```

<!--Load the AJAX API-->
<script type="text/javascript" src="https://www.google.com/jsapi"></script>
<script type="text/javascript">

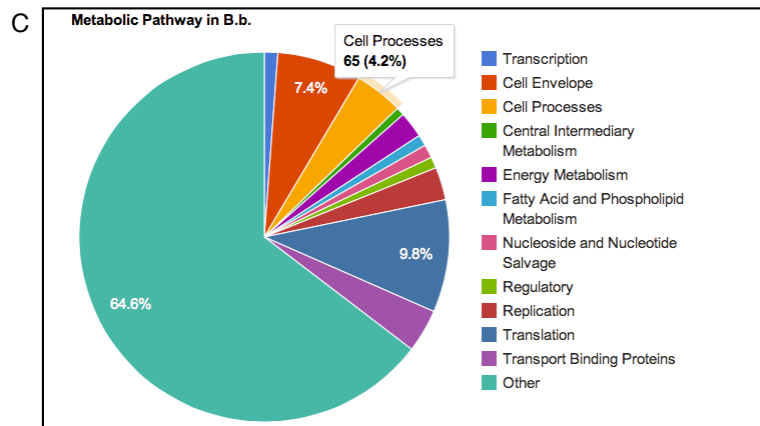
// Load the Visualization API and the piechart package.
google.load('visualization', '1.0', {'packages':['corechart']});

// Set a callback to run when the Google Visualization API is loaded.
google.setOnLoadCallback(drawChart);

// Callback that creates and populates a data table,
// instantiates the pie chart, passes in the data and
// draws it.
function drawChart() {

// Create the data table.
var data = new google.visualization.DataTable();

```



Supplemental Figure S10. Displaying Data on BTech. (A) This structure is a view, which is one of several built-in capabilities Drupal has to display data. It uses HTML and AJAX, and can either be customized with code or a graphical interface. (B) This is sample code for the Google® Charts API (Creative Commons License, used with permission). (C) Chart from BTech that displays the distribution of known gene coding types for *B. burgdorferi*.

LIST OF REFERENCES

1. Keller, A., et al., *New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing*. Nat Commun, 2012. **3**: p. 698.
2. Steere, A.C., et al., *Longitudinal assessment of the clinical and epidemiological features of Lyme disease in a defined population*. J Infect Dis, 1986. **154**(2): p. 295-300.
3. Shapiro, E.D., *Clinical practice. Lyme disease*. N Engl J Med, 2014. **370**(18): p. 1724-31.
4. Steere, A.C., et al., *Antibiotic therapy in Lyme disease*. Ann Intern Med, 1980. **93**(1): p. 1-8.
5. Steere, A.C. and S.E. Malawista, *Cases of Lyme disease in the United States: locations correlated with distribution of Ixodes dammini*. Ann Intern Med, 1979. **91**(5): p. 730-3.
6. Burgdorfer, W., et al., *Lyme disease-a tick-borne spirochetosis?* Science, 1982. **216**(4552): p. 1317-9.
7. Barbour, A.G., et al., *Antibodies of patients with Lyme disease to components of the Ixodes dammini spirochete*. J Clin Invest, 1983. **72**(2): p. 504-15.
8. Steere, A.C., et al., *The spirochetal etiology of Lyme disease*. N Engl J Med, 1983. **308**(13): p. 733-40.
9. Fraser, C.M., et al., *Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi*. Nature, 1997. **390**(6660): p. 580-6.
10. Piesman, J., et al., *Duration of tick attachment and Borrelia burgdorferi transmission*. J Clin Microbiol, 1987. **25**(3): p. 557-8.
11. Piesman, J., *Dynamics of Borrelia burgdorferi transmission by nymphal Ixodes dammini ticks*. J Infect Dis, 1993. **167**(5): p. 1082-5.
12. Feder, H.M., Jr., et al., *Early Lyme disease: a flu-like illness without erythema migrans*. Pediatrics, 1993. **91**(2): p. 456-9.
13. Reik Jr, L., W. Burgdorfer, and J.O. Donaldson, *Neurologic abnormalities in Lyme disease without erythema chronicum migrans*. Am J Med, 1986. **81**(1): p. 73-78.
14. Nadelman, R.B., et al., *The clinical spectrum of early Lyme borreliosis in patients with culture-confirmed erythema migrans*. Am J Med, 1996. **100**(5): p. 502-8.
15. Nadelman, R.B. and G.P. Wormser, *Lyme borreliosis*. Lancet, 1998. **352**(9127): p. 557-65.

16. Wormser, G.P., et al., *Practice guidelines for the treatment of Lyme disease. The Infectious Diseases Society of America*. Clin Infect Dis, 2000. **31 Suppl 1**: p. 1-14.
17. Feng, J., et al., *Identification of novel activity against Borrelia burgdorferi persists using an FDA approved drug library*. Emerg Microbes Infect, 2014. **3**: p. e49.
18. *CDC-Two Tier Tests-Lyme Disease*. November 15, 2011 8/9/2014]; Available from: http://www.cdc.gov/lyme/healthcare/clinician_twotier.html.
19. Rosa, P.A., K. Tilly, and P.E. Stewart, *The burgeoning molecular genetics of the Lyme disease spirochaete*. Nat Rev Microbiol, 2005. **3**(2): p. 129-43.
20. Barbieri, A.M., et al., *Borrelia burgdorferi sensu lato infecting ticks of the Ixodes ricinus complex in Uruguay: first report for the Southern Hemisphere*. Vector Borne Zoonotic Dis, 2013. **13**(3): p. 147-53.
21. Barbour, A.G. and S.F. Hayes, *Biology of Borrelia species*. Microbiol Rev, 1986. **50**(4): p. 381-400.
22. Charon, N.W. and S.F. Goldstein, *Genetics of motility and chemotaxis of a fascinating group of bacteria: the spirochetes*. Annu Rev Genet, 2002. **36**: p. 47-73.
23. Hovind-Hougen, K., *Ultrastructure of spirochetes isolated from Ixodes ricinus and Ixodes dammini*. Yale J Biol Med, 1984. **57**(4): p. 543-8.
24. Kenedy, M.R., T.R. Lenhart, and D.R. Akins, *The role of Borrelia burgdorferi outer surface proteins*. FEMS Immunol Med Microbiol, 2012. **66**(1): p. 1-19.
25. Motaleb, M.A., et al., *Borrelia burgdorferi periplasmic flagella have both skeletal and motility functions*. Proc Natl Acad Sci U S A, 2000. **97**(20): p. 10899-904.
26. Burgdorfer, W., et al., *Relationship of Borrelia burgdorferi to its arthropod vectors*. Scand J Infect Dis Suppl, 1991. **77**: p. 35-40.
27. Lane, R.S., *Susceptibility of the western fence lizard (Sceloporus occidentalis) to the Lyme borreliosis spirochete (Borrelia burgdorferi)*. Am J Trop Med Hyg, 1990. **42**(1): p. 75-82.
28. Lane, R.S., et al., *Refractoriness of the western fence lizard (Sceloporus occidentalis) to the Lyme disease group spirochete Borrelia bissettii*. J Parasitol, 2006. **92**(4): p. 691-6.
29. Magnarelli, L.A. and J.F. Anderson, *Ticks and biting insects infected with the etiologic agent of Lyme disease, Borrelia burgdorferi*. J Clin Microbiol, 1988. **26**(8): p. 1482-6.
30. Lane, R.S. and J.E. Loye, *Lyme disease in California: interrelationship of ixodid ticks (Acari), rodents, and Borrelia burgdorferi*. J Med Entomol, 1991. **28**(5): p. 719-25.
31. Lane, R.S., J. Piesman, and W. Burgdorfer, *Lyme borreliosis: relation of its causative agent to its vectors and hosts in North America and Europe*. Annu Rev Entomol, 1991. **36**: p. 587-609.
32. Radolf, J.D., et al., *Characterization of outer membranes isolated from Borrelia burgdorferi, the Lyme disease spirochete*. Infect Immun, 1995. **63**(6): p. 2154-63.

33. de Silva, A.M. and E. Fikrig, *Arthropod- and host-specific gene expression by Borrelia burgdorferi*. J Clin Invest, 1997. **99**(3): p. 377-9.
34. Schwan, T.G., et al., *Induction of an outer surface protein on Borrelia burgdorferi during tick feeding*. Proc Natl Acad Sci U S A, 1995. **92**(7): p. 2909-13.
35. Pal, U., et al., *A differential role for BB0365 in the persistence of Borrelia burgdorferi in mice and ticks*. J Infect Dis, 2008. **197**(1): p. 148-55.
36. Li, X., et al., *The Lyme disease agent Borrelia burgdorferi requires BB0690, a Dps homologue, to persist within ticks*. Mol Microbiol, 2007. **63**(3): p. 694-710.
37. Revel, A.T., et al., *bptA (bbe16) is essential for the persistence of the Lyme disease spirochete, Borrelia burgdorferi, in its natural tick vector*. Proc Natl Acad Sci U S A, 2005. **102**(19): p. 6972-7.
38. Carroll, J.A., C.F. Garon, and T.G. Schwan, *Effects of environmental pH on membrane proteins in Borrelia burgdorferi*. Infect Immun, 1999. **67**(7): p. 3181-7.
39. Casjens, S., et al., *A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi*. Mol Microbiol, 2000. **35**(3): p. 490-516.
40. Stewart, P.E., G. Chaconas, and P. Rosa, *Conservation of plasmid maintenance functions between linear and circular plasmids in Borrelia burgdorferi*. J Bacteriol, 2003. **185**(10): p. 3202-9.
41. Eggers, C.H., et al., *Identification of loci critical for replication and compatibility of a Borrelia burgdorferi cp32 plasmid and use of a cp32-based shuttle vector for the expression of fluorescent reporters in the Lyme disease spirochaete*. Mol Microbiol, 2002. **43**(2): p. 281-95.
42. Stewart, P.E., et al., *Isolation of a circular plasmid region sufficient for autonomous replication and transformation of infectious Borrelia burgdorferi*. Mol Microbiol, 2001. **39**(3): p. 714-21.
43. Simpson, W.J., C.F. Garon, and T.G. Schwan, *Analysis of supercoiled circular plasmids in infectious and non-infectious Borrelia burgdorferi*. Microb Pathog, 1990. **8**(2): p. 109-18.
44. Casjens, S., et al., *Telomeres of the linear chromosomes of Lyme disease spirochaetes: nucleotide sequence and possible exchange with linear plasmid telomeres*. Mol Microbiol, 1997. **26**(3): p. 581-96.
45. Kobryn, K. and G. Chaconas, *ResT, a telomere resolvase encoded by the Lyme disease spirochete*. Mol Cell, 2002. **9**(1): p. 195-201.
46. Chaconas, G., et al., *Telomere resolution in the Lyme disease spirochete*. EMBO J, 2001. **20**(12): p. 3229-37.
47. Iyer, R., et al., *Linear and circular plasmid content in Borrelia burgdorferi clinical isolates*. Infect Immun, 2003. **71**(7): p. 3699-706.
48. Grimm, D., et al., *Plasmid stability during in vitro propagation of Borrelia burgdorferi assessed at a clonal level*. Infect Immun, 2003. **71**(6): p. 3138-45.
49. Schwan, T.G., W. Burgdorfer, and C.F. Garon, *Changes in infectivity and plasmid profile of the Lyme disease spirochete, Borrelia burgdorferi, as a result of in vitro cultivation*. Infect Immun, 1988. **56**(8): p. 1831-6.

50. Guina, T. and D.B. Oliver, *Cloning and analysis of a Borrelia burgdorferi membrane-interactive protein exhibiting haemolytic activity*. Mol Microbiol, 1997. **24**(6): p. 1201-13.
51. Damman, C.J., et al., *Characterization of Borrelia burgdorferi BlyA and BlyB proteins: a prophage-encoded holin-like system*. J Bacteriol, 2000. **182**(23): p. 6791-7.
52. Eggers, C.H. and D.S. Samuels, *Molecular evidence for a new bacteriophage of Borrelia burgdorferi*. J Bacteriol, 1999. **181**(23): p. 7308-13.
53. Eggers, C.H., et al., *Bacteriophages of spirochetes*. J Mol Microbiol Biotechnol, 2000. **2**(4): p. 365-73.
54. Yang, X., et al., *Interdependence of environmental factors influencing reciprocal patterns of gene expression in virulent Borrelia burgdorferi*. Mol Microbiol, 2000. **37**(6): p. 1470-9.
55. Stevenson, B. and S.W. Barthold, *Expression and sequence of outer surface protein C among North American isolates of Borrelia burgdorferi*. FEMS Microbiol Lett, 1994. **124**(3): p. 367-72.
56. Ramamoorthy, R. and M.T. Philipp, *Differential expression of Borrelia burgdorferi proteins during growth in vitro*. Infect Immun, 1998. **66**(11): p. 5119-24.
57. Wang, P., et al., *Borrelia burgdorferi oxidative stress regulator BosR directly represses lipoproteins primarily expressed in the tick during mammalian infection*. Mol Microbiol, 2013. **89**(6): p. 1140-53.
58. Ouyang, Z., R.K. Deka, and M.V. Norgard, *BosR (BB0647) controls the RpoN-RpoS regulatory pathway and virulence expression in Borrelia burgdorferi by a novel DNA-binding mechanism*. PLoS Pathog, 2011. **7**(2): p. e1001272.
59. Ouyang, Z., J. Zhou, and M.V. Norgard, *Synthesis of RpoS is dependent on a putative enhancer binding protein Rrp2 in Borrelia burgdorferi*. PLoS One, 2014. **9**(5): p. e96917.
60. Groshong, A.M., et al., *Rrp2, a prokaryotic enhancer-like binding protein, is essential for viability of Borrelia burgdorferi*. J Bacteriol, 2012. **194**(13): p. 3336-42.
61. Smith, A.H., et al., *Evidence that RpoS (sigmaS) in Borrelia burgdorferi is controlled directly by RpoN (sigma54/sigmaN)*. J Bacteriol, 2007. **189**(5): p. 2139-44.
62. Hubner, A., et al., *Expression of Borrelia burgdorferi OspC and DbpA is controlled by a RpoN-RpoS regulatory pathway*. Proc Natl Acad Sci U S A, 2001. **98**(22): p. 12724-9.
63. Burtnick, M.N., et al., *Insights into the complex regulation of rpoS in Borrelia burgdorferi*. Mol Microbiol, 2007. **65**(2): p. 277-93.
64. Ouyang, Z., J.S. Blevins, and M.V. Norgard, *Transcriptional interplay among the regulators Rrp2, RpoN and RpoS in Borrelia burgdorferi*. Microbiology, 2008. **154**(Pt 9): p. 2641-58.
65. Caimano, M.J., et al., *RpoS is not central to the general stress response in Borrelia burgdorferi but does control expression of one or more essential virulence determinants*. Infect Immun, 2004. **72**(11): p. 6433-45.

66. Hartmann, K., et al., *Functional characterization of BbCRASP-2, a distinct outer membrane protein of Borrelia burgdorferi that binds host complement regulators factor H and FHL-1*. Mol Microbiol, 2006. **61**(5): p. 1220-36.
67. Bhattacharjee, A., et al., *Structural Basis for Complement Evasion by Lyme Disease Pathogen Borrelia burgdorferi*. J Biol Chem, 2013. **288**(26): p. 18685-95.
68. Embers, M.E., R. Ramamoorthy, and M.T. Philipp, *Survival strategies of Borrelia burgdorferi, the etiologic agent of Lyme disease*. Microbes Infect, 2004. **6**(3): p. 312-8.
69. von Lackum, K., et al., *Borrelia burgdorferi regulates expression of complement regulator-acquiring surface protein 1 during the mammal-tick infection cycle*. Infect Immun, 2005. **73**(11): p. 7398-405.
70. Tilly, K., A. Bestor, and P.A. Rosa, *Lipoprotein succession in Borrelia burgdorferi: similar but distinct roles for OspC and VlsE at different stages of mammalian infection*. Mol Microbiol, 2013. **89**(2): p. 216-27.
71. Ohnishi, J., et al., *Genetic variation at the vlsE locus of Borrelia burgdorferi within ticks and mice over the course of a single transmission cycle*. J Bacteriol, 2003. **185**(15): p. 4432-41.
72. Zhang, J.R., et al., *Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes*. Cell, 1997. **89**(2): p. 275-85.
73. Thattai, M. and A. van Oudenaarden, *Stochastic gene expression in fluctuating environments*. Genetics, 2004. **167**(1): p. 523-30.
74. Baranyi, J., *Comparison of Stochastic and Deterministic Concepts of Bacterial Lag*. J Theor Biol, 1998. **192**(3): p. 403-408.
75. Lewis, K., *Persister cells*. Annu Rev Microbiol, 2010. **64**: p. 357-72.
76. Bockenstedt, L.K., et al., *What ticks do under your skin: two-photon intravital imaging of Ixodes scapularis feeding in the presence of the Lyme disease spirochete*. Yale J Biol Med, 2014. **87**(1): p. 3-13.
77. Tilly, K., P.A. Rosa, and P.E. Stewart, *Biology of infection with Borrelia burgdorferi*. Infect Dis Clin North Am, 2008. **22**(2): p. 217-34, v.
78. Embers, M.E., et al., *Persistence of Borrelia burgdorferi in rhesus macaques following antibiotic treatment of disseminated infection*. PLoS One, 2012. **7**(1): p. e29914.
79. Roberts, E.D., et al., *Pathogenesis of Lyme neuroborreliosis in the rhesus monkey: the early disseminated and chronic phases of disease in the peripheral nervous system*. J Infect Dis, 1998. **178**(3): p. 722-32.
80. Philipp, M.T. and B.J. Johnson, *Animal models of Lyme disease: pathogenesis and immunoprophylaxis*. Trends Microbiol, 1994. **2**(11): p. 431-7.
81. Philipp, M.T., et al., *Early and early disseminated phases of Lyme disease in the rhesus monkey: a model for infection in humans*. Infect Immun, 1993. **61**(7): p. 3047-59.
82. Barthold, S.W., et al., *Lyme borreliosis in selected strains and ages of laboratory mice*. J Infect Dis, 1990. **162**(1): p. 133-8.
83. Elias, A.F., et al., *Clonal polymorphism of Borrelia burgdorferi strain B31 MI: implications for mutagenesis in an infectious strain background*. Infect Immun, 2002. **70**(4): p. 2139-50.

84. Kelly, R., *Cultivation of Borrelia hermsi*. Science, 1971. **173**(3995): p. 443-4.
85. Stoenner, H.G., T. Dodd, and C. Larsen, *Antigenic variation of Borrelia hermsii*. J Exp Med, 1982. **156**(5): p. 1297-311.
86. Barbour, A.G., *Isolation and cultivation of Lyme disease spirochetes*. Yale J Biol Med, 1984. **57**(4): p. 521-5.
87. Alban, P.S., P.W. Johnson, and D.R. Nelson, *Serum-starvation-induced changes in protein synthesis and morphology of Borrelia burgdorferi*. Microbiology, 2000. **146** (Pt 1): p. 119-27.
88. Nocton, J.J., et al., *Detection of Borrelia burgdorferi DNA by polymerase chain reaction in synovial fluid from patients with Lyme arthritis*. N Engl J Med, 1994. **330**(4): p. 229-34.
89. Resources For Clinicians. 2014; Available from: <http://www.cdc.gov/lyme/healthcare/clinicians.html>.
90. Pahl, A., et al., *Quantitative detection of Borrelia burgdorferi by real-time PCR*. J Clin Microbiol, 1999. **37**(6): p. 1958-63.
91. Schmidt, B.L., et al., *Detection of Borrelia burgdorferi DNA by polymerase chain reaction in the urine and breast milk of patients with Lyme borreliosis*. Diagn Microbiol Infect Dis, 1995. **21**(3): p. 121-8.
92. Iyer, R., et al., *Detection of Borrelia burgdorferi nucleic acids after antibiotic treatment does not confirm viability*. J Clin Microbiol, 2013. **51**(3): p. 857-62.
93. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nat Rev Genet, 2009. **10**(1): p. 57-63.
94. Nagalakshmi, U., et al., *The transcriptional landscape of the yeast genome defined by RNA sequencing*. Science, 2008. **320**(5881): p. 1344-9.
95. Mortazavi, A., et al., *Mapping and quantifying mammalian transcriptomes by RNA-Seq*. Nat Methods, 2008. **5**(7): p. 621-8.
96. Smoot, L.M., et al., *Global differential gene expression in response to growth temperature alteration in group A Streptococcus*. Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10416-21.
97. Meyerson, M., S. Gabriel, and G. Getz, *Advances in understanding cancer genomes through second-generation sequencing*. Nat Rev Genet, 2010. **11**(10): p. 685-96.
98. Croucher, N.J. and N.R. Thomson, *Studying bacterial transcriptomes using RNA-seq*. Curr Opin Microbiol, 2010. **13**(5): p. 619-24.
99. Anders, S. and W. Huber, *Differential expression analysis for sequence count data*. Genome Biol, 2010. **11**(10): p. R106.
100. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data*. Bioinformatics, 2010. **26**(1): p. 139-40.
101. Hardcastle, T.J. and K.A. Kelly, *baySeq: empirical Bayesian methods for identifying differential expression in sequence count data*. BMC Bioinformatics, 2010. **11**: p. 422.
102. Lyme Disease Data. [cited 2014 10/21/2014]; Available from: <http://www.cdc.gov/lyme/stats/index.html>.
103. Aucott, J., et al., *Diagnostic challenges of early Lyme disease: lessons from a community case series*. BMC Infect Dis, 2009. **9**: p. 79.

104. Feder, H.M., Jr., et al., *A critical appraisal of "chronic Lyme disease"*. N Engl J Med, 2007. **357**(14): p. 1422-30.
105. Hodzic, E., et al., *Resurgence of persisting non-cultivable Borrelia burgdorferi following antibiotic treatment in mice*. PLoS One, 2014. **9**(1): p. e86907.
106. Steere, A.C., J. Coburn, and L. Glickstein, *The emergence of Lyme disease*. J Clin Invest, 2004. **113**(8): p. 1093-101.
107. Bockenstedt, L.K., et al., *Spirochete antigens persist near cartilage after murine Lyme borreliosis therapy*. J Clin Invest, 2012. **122**(7): p. 2652-60.
108. Ohnishi, M., et al., *Is Neisseria gonorrhoeae initiating a future era of untreatable gonorrhea?: detailed characterization of the first strain with high-level resistance to ceftriaxone*. Antimicrob Agents Chemother, 2011. **55**(7): p. 3538-45.
109. Chopra, I. and M. Roberts, *Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance*. Microbiol Mol Biol Rev, 2001. **65**(2): p. 232-60 ; second page, table of contents.
110. Yotsuji, A., et al., *Mechanism of action of cephalosporins and resistance caused by decreased affinity for penicillin-binding proteins in Bacteroides fragilis*. Antimicrob Agents Chemother, 1988. **32**(12): p. 1848-53.
111. Dean, C.R., et al., *Efflux-mediated resistance to tigecycline (GAR-936) in Pseudomonas aeruginosa PAO1*. Antimicrob Agents Chemother, 2003. **47**(3): p. 972-8.
112. Kumarasamy, K.K., et al., *Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study*. Lancet Infect Dis, 2010. **10**(9): p. 597-602.
113. Bigger, J.W., *Treatment of staphylococcal infections with penicillin*. Lancet, 1944. **244**: p. 497-500.
114. Lewis, K., *Persister cells, dormancy and infectious disease*. Nat Rev Microbiol, 2007. **5**(1): p. 48-56.
115. Dorr, T., M. Vulic, and K. Lewis, *Ciprofloxacin causes persister formation by inducing the TisB toxin in Escherichia coli*. PLoS Biol, 2010. **8**(2): p. e1000317.
116. Mulcahy, L.R., et al., *Emergence of Pseudomonas aeruginosa strains producing high levels of persister cells in patients with cystic fibrosis*. J Bacteriol, 2010. **192**(23): p. 6191-9.
117. Hodzic, E., et al., *Persistence of Borrelia burgdorferi following antibiotic treatment in mice*. Antimicrob Agents Chemother, 2008. **52**(5): p. 1728-36.
118. Patton, T.G., et al., *Borrelia burgdorferi bba66 gene inactivation results in attenuated mouse infection by tick transmission*. Infect Immun, 2013. **81**(7): p. 2488-98.
119. Barthold, S.W., et al., *Ineffectiveness of tigecycline against persistent Borrelia burgdorferi*. Antimicrob Agents Chemother, 2010. **54**(2): p. 643-51.
120. Embers, M.E. and S.W. Barthold, *Borrelia burgdorferi Persistence Post-antibiotic Treatment*, in *The Pathogenic Spirochetes: strategies for evasion of host immunity and persistence* 2012, Springer. p. 229-257.
121. Lewis, K., *Persister cells: molecular mechanisms related to antibiotic tolerance*. Handb Exp Pharmacol, 2012(211): p. 121-33.

122. Purser, J.E. and S.J. Norris, *Correlation between plasmid content and infectivity in Borrelia burgdorferi*. Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13865-70.
123. Wodecka, B., *flaB gene as a molecular marker for distinct identification of Borrelia species in environmental samples by the PCR-restriction fragment length polymorphism method*. Appl Environ Microbiol, 2011. **77**(19): p. 7088-92.
124. Signorino, G., et al., *Identification of OppA2 linear epitopes as serodiagnostic markers for Lyme disease*. Clin Vaccine Immunol, 2014. **21**(5): p. 704-11.
125. *Borrelia burgdorferi B31 RefSeq Genome*, 2010: <http://www.ncbi.nlm.nih.gov/bioproject/PRJNA57581>.
126. Knowles, D.J., et al., *The bacterial ribosome, a promising focus for structure-based drug design*. Curr Opin Pharmacol, 2002. **2**(5): p. 501-6.
127. Davidson, A.L., et al., *Structure, function, and evolution of bacterial ATP-binding cassette systems*. Microbiol Mol Biol Rev, 2008. **72**(2): p. 317-64, table of contents.
128. Stevenson, B., T.G. Schwan, and P.A. Rosa, *Temperature-related differential expression of antigens in the Lyme disease spirochete, Borrelia burgdorferi*. Infect Immun, 1995. **63**(11): p. 4535-9.
129. Phillips, I., et al., *Induction of the SOS response by new 4-quinolones*. J Antimicrob Chemother, 1987. **20**(5): p. 631-8.
130. Lewis, K., *Multidrug tolerance of biofilms and persister cells*. Curr Top Microbiol Immunol, 2008. **322**: p. 107-31.
131. Balaban, N.Q., et al., *Bacterial persistence as a phenotypic switch*. Science, 2004. **305**(5690): p. 1622-5.
132. Codling, E.A., M.J. Plank, and S. Benhamou, *Random walk models in biology*. J R Soc Interface, 2008. **5**(25): p. 813-34.
133. Rainey, P.B., et al., *The evolutionary emergence of stochastic phenotype switching in bacteria*. Microb Cell Fact, 2011. **10 Suppl 1**: p. S14.
134. Nguyen, D., et al., *Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria*. Science, 2011. **334**(6058): p. 982-6.
135. Wu, Y., et al., *Role of oxidative stress in persister tolerance*. Antimicrob Agents Chemother, 2012. **56**(9): p. 4922-6.
136. Balaban, N.Q., et al., *A problem of persistence: still more questions than answers?* Nat Rev Microbiol, 2013. **11**(8): p. 587-91.
137. Avery, S.V., *Microbial cell individuality and the underlying sources of heterogeneity*. Nat Rev Microbiol, 2006. **4**(8): p. 577-87.
138. Partek, I., *White Paper: RNAseq Methods*, Partek, Editor 2013.
139. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
140. Kane, M.D., et al., *Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays*. Nucleic Acids Res, 2000. **28**(22): p. 4552-7.
141. Shendure, J. and H. Ji, *Next-generation DNA sequencing*. Nat Biotechnol, 2008. **26**(10): p. 1135-45.
142. Metzker, M.L., *Sequencing technologies - the next generation*. Nat Rev Genet, 2010. **11**(1): p. 31-46.

143. Harris, T.D., et al., *Single-molecule DNA sequencing of a viral genome*. Science, 2008. **320**(5872): p. 106-9.
144. Adessi, C., et al., *Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms*. Nucleic Acids Res, 2000. **28**(20): p. E87.
145. Bentley, D.R., et al., *Accurate whole human genome sequencing using reversible terminator chemistry*. Nature, 2008. **456**(7218): p. 53-9.
146. Kozarewa, I., et al., *Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes*. Nat Methods, 2009. **6**(4): p. 291-5.
147. Merriman, B. and J.M. Rothberg, *Progress in ion torrent semiconductor chip based sequencing*. Electrophoresis, 2012. **33**(23): p. 3397-417.
148. Quail, M.A., et al., *A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers*. BMC Genomics, 2012. **13**: p. 341.
149. Rothberg, J.M., et al., *An integrated semiconductor device enabling non-optical genome sequencing*. Nature, 2011. **475**(7356): p. 348-52.
150. Schadt, E.E., S. Turner, and A. Kasarskis, *A window into third-generation sequencing*. Hum Mol Genet, 2010. **19**(R2): p. R227-40.
151. Turcatti, G., et al., *A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis*. Nucleic Acids Res, 2008. **36**(4): p. e25.
152. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors*. Nature, 2005. **437**(7057): p. 376-80.
153. Levene, M.J., et al., *Zero-mode waveguides for single-molecule analysis at high concentrations*. Science, 2003. **299**(5607): p. 682-6.
154. Travers, K.J., et al., *A flexible and efficient template format for circular consensus sequencing and SNP detection*. Nucleic Acids Res, 2010. **38**(15): p. e159.
155. Ewing, B. and P. Green, *Base-calling of automated sequencer traces using phred. II. Error probabilities*. Genome Res, 1998. **8**(3): p. 186-94.
156. Li, H., J. Ruan, and R. Durbin, *Mapping short DNA sequencing reads and calling variants using mapping quality scores*. Genome Res, 2008. **18**(11): p. 1851-8.
157. Mamanova, L., et al., *FRT-seq: amplification-free, strand-specific transcriptome sequencing*. Nat Methods, 2010. **7**(2): p. 130-2.
158. Haas, B.J., et al., *How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes?* BMC Genomics, 2012. **13**: p. 734.
159. Hawkins, T.L., et al., *DNA purification and isolation using a solid-phase*. Nucleic Acids Res, 1994. **22**(21): p. 4543-4.
160. Elkin, C., et al., *Magnetic bead purification of labeled DNA fragments for high-throughput capillary electrophoresis sequencing*. Biotechniques, 2002. **32**(6): p. 1296, 1298-1300, 1302.
161. Technologies, L., *Life Technologies Ion Total RNA-Seq Kit v2*, in *Life Technologies* 2013.
162. Zhuang, F., et al., *Structural bias in T4 RNA ligase-mediated 3'-adapter ligation*. Nucleic Acids Res, 2012. **40**(7): p. e54.

163. *Protocol from Invitrogen® for Ribominus Transcriptome Isolation Kit Bacteria and Yeast.*
164. Cormen, T.H., Leiserson, Charles E., Rivest, Ronald L., Stein, Clifford, *Introduction to Algorithms*. 3rd edc2009: MIT Press. 1292.
165. Jones, N.C., Pevzner, Pavel A., *An Introduction to Bioinformatics Algorithms*. 1st edc2004: MIT Press.
166. Burrows, M., Wheeler DJ, *A block-sorting lossless data compression algorithm*. 1994.
167. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform*. *Bioinformatics*, 2009. **25**(14): p. 1754-60.
168. Li, H. and N. Homer, *A survey of sequence alignment algorithms for next-generation sequencing*. *Brief Bioinform*, 2010. **11**(5): p. 473-83.
169. Smith, T.F. and M.S. Waterman, *Identification of common molecular subsequences*. *J Mol Biol*, 1981. **147**(1): p. 195-7.
170. Google, I. *Google Inc.* [cited 2014; Creative Commons License, Used With Permission]. Available from: <https://developers.google.com/chart/>.

BIOGRAPHY

John Russell Caskey was born in Baton Rouge, Louisiana. After studying History and Theatre, he graduated with a B.A. from The University of the South at Sewanee, in Sewanee TN, in 2004, and completed a B.S. at Florida Atlantic University in 2009. He entered the Tulane Biomedical Sciences Ph.D. program, in the fall of 2010, and joined the Embers Lab at the Tulane National Primate Research Center. To help his analysis of *B. burgdorferi*, he wrote and published a cell counting App on the Apple® App Store, and created a searchable database of the *B. burgdorferi* genome. He also does statistical analysis for colleagues, and is a contributing author in the publication "Immunomodulatory effects of tick saliva on dermal cells exposed to *Borrelia burgdorferi*, the agent of Lyme disease". He currently has a paper in publication that uses part of his doctoral research, "Persister Development in *B. burgdorferi* Populations", and was a contributing author in the paper, "Septic Arthritis Due to *Moraxella osloensis* in a Rhesus Macaque (*Macaca mulatta*)" (Comparative Medicine 2013).