Identification and Characterization of Molecular Correlates of Disease in Invasive *Streptococcus* and *Staphylococcus* Infection Models

by

Richard William Davis IV

A dissertation submitted to the Graduate Faculty of
Auburn University
In partial fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 1st, 2015

Approved by

Peter Panizzi, Chair, Assistant Professor of Drug Discovery and Development
Mark R. Liles, Associate Professor of Biology
Jianzhong Shen, Associate Professor of Drug Discovery and Development
Jack Deruiter, Professor of Drug Development and Discovery

Abstract

Invasion into the host vasculature marks a dangerous, and often times fatal, complication of bacterial infections. One such severely invasive bacterium, *Streptococcus pyogenes*, is estimated to affect 18 million people globally, resulting in 517,000 deaths per year. In order to invade the host, *S. pyogenes* and others rely on an intricate and precisely-timed network of virulence factors and host response. To properly characterize this interplay, as well as the redundancies inherent in each, there is a need for technology which allows real-time characterization of bacterial factors and host immune response during invasive infections. In this dissertation, we will cover the use of two such technologies, namely molecular imaging and next-generation sequencing, towards the advancement of characterizing this essential interplay. Although typically considered as two separate disciplines, this dissertation will demonstrate the means by which these technologies work in tandem to produce new information as to the timing and composition of the infectious system. Within, we show that *S. pyogenes* is able to disseminate in host tissue without the use of streptokinase through the use of the genetically engineered light-producing pathogen Xen20. We also show that the luminescence production of this strain is dependent on both genetic regulation by the *iol* operon, and by catabolite repression of bacterial central metabolism. By exploiting this central metabolism-based indicator of cell physiology, we were able to design experiments which determine the mechanism of antibiosis of known compounds, indicating a potential for future high-throughput characterizations of lead targets. We also outline the automated construction and characterization of the *Staphylococcus aureus* Tager 104 genome, a potential "missing link" in the study of Staphylococcal pathogenesis.

Acknowledgements

I am grateful to many for the completion of this dissertation. First, to my advisor, Peter Panizzi, and to the members of my committee for their time and dedication. I am thankful also to Rajesh Amin, who has on multiple accounts supported this work and my dissertation studies. I am thankful for my friends and family for their support, especially my mother, Alicia, my father, Rick, and my sister, Haley. I am humbled and thankful for the constant love and support of MariAnne Sullivan throughout the course of this research and writing. I am thankful for the support my friends, especially Andrew Brannen, Jiansheng Huang, Rob and J.D. Yantz, and Carlos Gonzalez. Thanks to the Dave Young and the Alabama Supercomputer Authority for use of the Alabama Supercomputer system. Thanks to the University of Florida Interdisciplinary Center for Biotechnology Research for the Pacific Biosciences (PacBio) sequencing runs. Thanks to David Mead and Scott Monsma of Lucigen Corporation for construction and validation of the Lucigen NxSeq libraries. Thanks to Auburn University and the National Institute of Health for funding the research within. Thanks also to co-authors and collaborators on these research projects, including: Heather Eggleston, Jahangir Hossain, Matthias Nahrendorf, and Paul Bock.

Table of Contents

Abstract ii

Acknowledgements iii

List of Tables ix

List of Illustrations x

1. Combatting Invasive Infections Using Next Generation Technologies: An Overview 1

Gram Positive Invasive Infections 1

Invasive Infections Caused by Streptococcus pyogenes 3

Staphylococcus aureus Invasion and Survival Systemically 5

Staphylococcus aureus Endocarditis: A Consequence of Invasive Infections 6

Staphylococcal Genomics 8

Systems Biology Approaches to Characterizing Invasive Diseases 9

2. Streptokinase-Deficient Dissemination of *Streptococcus pyogenes in vivo* 15

Introduction 15

Materials and Methods 19

Animals 19

Plasminogen Activation Assay 20

Preparation of Bacterial Injections 20

Animal Procedures 21

Histology and Ex Vivo Imaging 21

Glucose Dependence Assay 22

Results 23

Activation of Human Plasminogen by S. pyogenes Xen20 and ALAB49 Supernatants 23

Genomic Analysis of S. pyogenes Plasminogen Activation System 25

Diminished Bioluminescence at the Site of Infection for Xen10 and Xen20 strains 26

Gram Staining and CFU Determination for Pathogens in Distal Organs and the Skin 28

Presence of Exogenous Glucose Modulates Xen20 Bioluminescence in a Concentration Dependent Manner 32

Discussion 35

3. Dependence of Bioluminescence Expression on Central Metabolites in *S. pyogenes* 39

Introduction 39

Materials and Methods 41

Reagents, Chemicals, and Strains 41

Bioluminescence Plate Assay 41

Bioluminescence Kinetic Assays 41

Luciferase Protein Assay 43

Oxamate Inhibition Assay 43

Statistical Analysis 44

Results 44

D-Glucose Inhibits Bioluminescence from Xen20 44

The lux Operon Is Regulated by Inositol and Glucose Levels 46

Substrate Sources for Streptococcus pyogenes Bioluminescence 48

Dependence of Luminescence on Glucose Homeostasis 52

Dependence of Energetics on Luminescence Expression 58

Inhibition of L-lactate dehydrogenase 60

In vitro Model of Tissue Luminescence 62

Discussion 64

Application Note: Characterizing Antibiotic Mechanisms Using Bioluminescence Imaging. 68

4. Genomic and Pathogenic Characterization of an Ancestral *Staphylococcus aureus* Strain Potent in Systemic Models of Infection 72

Introduction 72

Materials and Methods 74

Mouse Model of Endocarditis. 74

Fluorescence Molecular Tomography fused to Computer Tomography (FMT-CT) Detection of Endocarditis. 75
Mouse Model of Systemic Infection. 75
Histology. 76
Tager 104 Genome Construction. 76
Construction Validation and Analysis. 77
MLST Analysis. 77
Proteomic Analysis of Tager 104 Genome. 78
Whole-Genome Phylogenetic Analysis. 79
MGE-encoded Gene Determinations. 81
Antibiotic Susceptibility Testing. 81
Results 82
Tager 104 Forms Potent Endocarditis in vivo 82
S. aureus Tager 104 Abscess Formation in Naïve Tissue 84
Automated Construction of S. aureus Tager 104 Genome. 86
Repeat Regions Created Difficulty in Automated Closure 88
Proteomic Analysis of S. aureus Reference Sequences 90
Tager 104 is an early-branching ancestor to modern clinical strains. 92

Tager 104 Contains a Limited Set of MGE-encoded Virulence Factors Targeted at Host Immunity 94

Tager 104 Shows Intermediate Clinical Adaptations 100

S. aureus Tager 104 Resistance Profile 102

Discussion 104

5. Conclusions, Discussion, and Future Directions 109

References 111

List of Tables

Table 1.1: Aims, Hypothesis, and Outline of This Dissertation…….14

Table 3.1: Two Way Anova Results for Substrate-Dependent Lumiescence……..…….….57

Table 4.1: Bacterial Strains Used In This Study……...….80

Table 4.2: Genomic Island Content in *S. aureus* Tager 104……....96

Table 4.3: Prophage Content in *S. aureus* Tager 104…….98

List of Illustrations

Figure 1.1: National Estimates of Invasive Disease by Bacterial Species. 2

Figure 1.2: Cyclical Systeomic Approaches Using Next-Generation Technologies 13

Figure 2.1: Human Plasminogen Activation by the Streptococcal Secretomes. 24

Figure 2.2: Time course of Streptococcal Infection Monitored by Bioluminescence Imaging. ... 27

Figure 2.3: Histological analysis of Streptococcal strains reveals widespread dissemination. 29

Figure 2.4: Evidence of bioluminescent colonies retrieved from *in vivo* infection. 31

Figure 2.5: Modulation of bioluminescence as a function of glucose concentration in a light producing *S. pyogenes*. 34

Figure 3.1: Glucose Effect on Bioluminescent Production of *S. pyogenes* Xen 20. 45

Figure 3.2: *Streptococcus pyogenes* Xen20 luminescence is inositol-dependent 47

Figure 3.3: *De novo* synthesis of acyl donors via central carbon metabolism in *Streptococcus*. . 49

Figure 3.4: KEGG Genome Analysis of *S. pyogenes* and *S. aureus* β-oxidation Pathways. 51

Figure 3.5: Evidence of bioluminescent colonies retrieved from *in vivo* infection. 53

Figure 3.6: Increased luminescence production by Gram-positive pathogen incubated in M9 medium supplemented with casein hydrolysate and yeast extract. 55

Figure 3.7: Partial recovery of luminescence signal upon addition of D-glucose and ATP or AMP, but not ATP or AMP alone 59

Figure 3.8: Oxamate inhibition does not recover luminescence signal. 61

Figure 3.9: Effect of blood supplementation of biolumenescence signal propagation by *S. pyogenes* Xen20. 63

Figure 3.10: Medium Dependence of Lumincescence on Antimicrobial Differentation. 71

Figure 4.1: *S. aureus* Tager 104 Endocarditis..... 83

Figure 4.2: *S. aureus* Tager 104 Bacteremia Leading to Multi-organ Septic Foci. 85

Figure 4.3: Construction of the Tager 104 Genome 87

Figure 4.4: Graphical Depiction of Tager 104 Assembly..... 89

Figure 4.5: Proteomic Analysis of *S. aureus* Reference Sequences 91

Figure 4.6: Whole-Genome Phylogenetic Analysis of *S. aureus* Reference Strains 93

Figure 4.7: Mobile Genetic Element-Encoded Virulence in *S. aureus* Tager 104. 95

Figure 4.8: Tager 104 νSaβ Genomic Island Shows Intermediate Clinical Adaptation. 101

Figure 4.9: *S. aureus* Tager 104 Susceptibility to Antibiotic Therapy Demonstrates Predation of Resistance Development. 103

1. Combatting Invasive Infections Using Next Generation Technologies: An Overview

Gram Positive Invasive Infections

Invasive infections occur when the bacterium escapes from an initial site of infection, such as the skin, and penetrates into the bloodstream. In the year 2012, the Active Bacterial Core Surveillance for the Centers for Disease Control determined the national estimate of invasive diseases to be a combined 83,994 cases [1]. Increasing this concern is the recent estimate by the Centers for Disease Control that more than two million people are sickened every year with antibiotic-resistant infections, causing an estimated 23,000 or more deaths [2]. Figure 1.1 demonstrates the national estimate of invasive disease, based on data collected by the 2012 Active Bacterial Core Surveillance [1]. This information is provided in the format of number of cases (shown in blue bars) and number of mortalities (shown in orange bars) per 100,000 members of the population. As demonstrated, Gram positive bacteria compose the vast majority of these cases, especially *Streptococcus* and methicillin resistant *Staphylococcus aureus* (MRSA). Based on this observation, the focus of this dissertation will be the characterization of molecular correlates that these bacterial species possess through holistic, systems biology approaches.

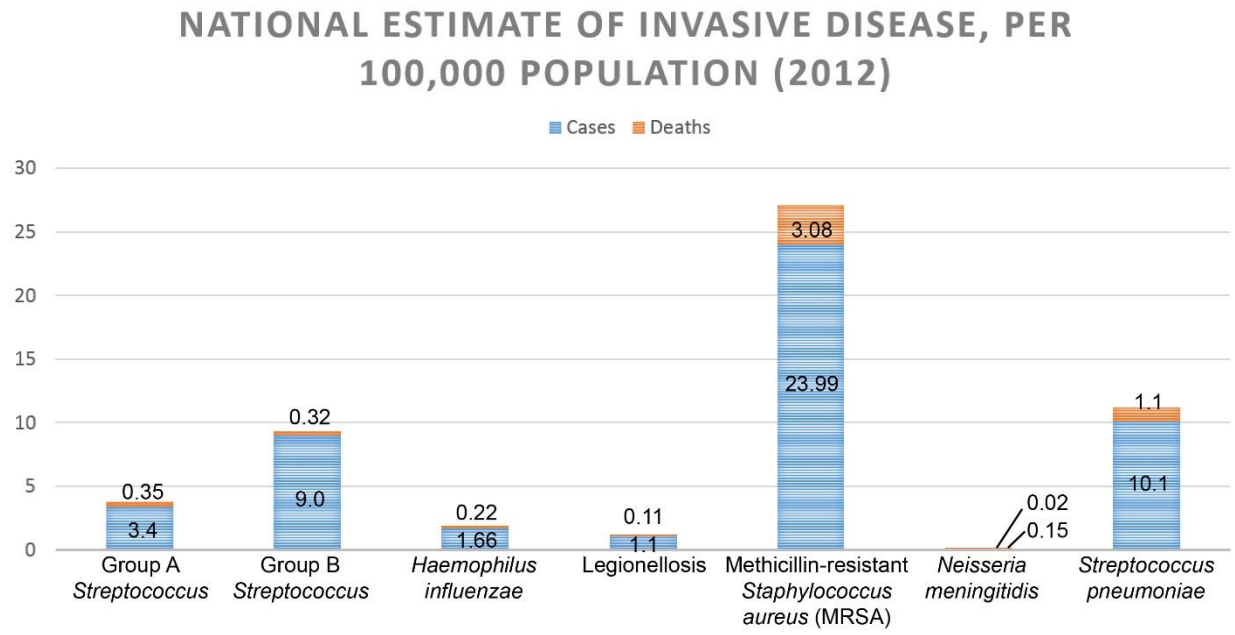

Figure 1.1: National Estimates of Invasive Disease by Bacterial Species.

National estimates (per 100,000 population) are shown as the number of cases (*blue rectangles*) and mortalities (*orange rectangles*) for bacterial species. Information adapted from the Active Bacterial Core surveillance (ABCs) report for 2012.

Invasive Infections Caused by Streptococcus pyogenes

Dangerous group A Streptococci (GAS) infections are estimated to cause a global impact of 18 million people, resulting in 517,000 deaths per year [3]. According to the 2013 Active Bacterial Core Surveillance (ABCs), GAS caused approximately 9,000-11,500 cases in the United States, resulting in 1,000-1,800 deaths [1] One particularly dangerous scenario in these infections is the penetration of *S. pyogenes* from localized infections to distal tissue, a process known as dissemination. Once the bacterium has penetrated through the epithelial layer, it can produce an array of diseases such as bacteremia, cellulitis, or necrotizing fasciitis. An extensive review of GAS infections can be found elsewhere [4]. These diseases can be further intensified by the onset of a condition known as streptococcal toxic shock-like syndrome (STSS). In the clinical setting, 19% of people who develop invasive diseases die within 7 days, and this number increases to 44% in the presence of STSS [5].

Streptococcal disease begins with the bacterium adhering to the host tissue. This adherence is strengthened by lectin-carbohydrate and protein-protein interactions, such as pili, M protein, and extracellular matrix-binding proteins [6, 7]. This M-protein is also used to serotype strains, and M serotype seems to correlate very strongly with invasivity and pathogenesis [4]. After adhering and penetrating the skin layer, GAS initiate a wide range of answers to host response. First, GAS express several virulence factors to evade the host immune response. The hyaluronic acid capsule of *S. pyogenes* is a polymer of glucuronic-β-1,3-*N*-acetylglucosamine, and provides resistance to opsonization and phagocytosis [8, 9]. More directly, interleukin-8 protease (SpyCEP) [10] and immunoglobulin G endopeptidase (IdeS) [11] cleave IL-8 and the heavy chain of IgG, respectively, preventing neutrophil and Fc-mediated opsonophagocytosis. To prevent complement-mediated killing, GAS release the Streptococcal inhibitor of complement (SIC), which prevents the

interaction between C5b67 and host cell membranes, eliminating the membrane attack complex [12, 13].

Several GAS proteins have been associated with dissemination. These include the Streptococcal secreted esterase (Sse), the cysteine protease SpeB, and streptokinase (SK) [4]. In murine models, deletion of *sse* (Δ*sse*) was associated with a significant decrease in bacterial dissemination to the blood and spleen, as well as a decreased ability to survive in the subcutis [14]. SpeB, a cysteine protease expressed highly at the site of infection in necrotizing fasciitis, cleaves targets such as ECM, cytokine precursors, immunoglobulins, and antimicrobial peptides [15]. As it is also known to cleave bacterial virulence factors, it is believed that the expression of this protein occurs in early stages of infection, and downregulation late in the infection may counter a mounting hyperinflammatory response [16]. In support of this theory, deletion of the *speB* gene in *S. pyogenes* M1T1, a highly invasive clonal type, demonstrated a significant 75-fold increase in plasmin activity on the surface of the cell, driven by the streptokinase/plasminogen interface [15]. This accumulation of plasmin on the cell surface has been associated with invasive infections [17]. Streptokinase is a non-enzymatic activator of plasminogen. This activation can happen in one of two ways: the direct pathway, in which plasminogen binds directly to receptors such as the plasminogen-binding group A streptococcal M-like protein (PAM, PAM-related protein (PrP), streptococcal enolase (Eno) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH); and the indirect pathway, in which formation of a streptokinase-plasminogen-fibrinogen complex binds to plasminogen or fibrinogen receptors [4]. This streptokinase can be clustered into three major groups, cluster 1, cluster 2a, and cluster 2b, and will be discussed further in Chapter 2. Interestingly, mouse plasminogen seems to have no effect with streptokinase. Mice injected subcutaneously with a streptokinase-deficient (Δ*ska*) *S. pyogenes* strain, UMAA2616, showed no

significant difference in skin lesion area or percent survival [18]. Co-administration of human plasminogen with this strain increased the mean skin lesion area and decreased the percent survival [18]. Similarly, humanized plasminogen transgenic mice were injected with GAS strain 5448 and showed a 50% survival rate. This rate dropped to 10% when strain NZ131 streptokinase was cloned into 5448, and dropped to 10% in strains which carried either ALAB49 streptokinase or a genetic deletion [19]. Therefore, in this dissertation, we will cover the streptokinase-independent dissemination of *S. pyogenes* (Chapter 2) using a mouse model, which is naturally devoid of the streptokinase-plasminogen interface.

Staphylococcus aureus Invasion and Survival Systemically

Staphylococcus aureus is a robust bacterium, and the leading cause of nosocomial infections in the United States. In 2009, a study produced by Harvard Medical School indicated the total burden of cost for healthcare-associated infections alone to be $9.78 billion, with MRSA-based surgical or central line-associated infections contributing $1.38 billion alone [20]. Other estimates are much higher for these infections, predicting values as high as $45 billion (http://www.cdc.gov/HAI/pdfs/hai/Scott_CostPaper.pdf).

Bacteremia is a serious outcome of *Staphylococcus aureus* infections, as it can lead to complications such as infective endocarditis [21] and toxic shock syndrome (TSS) [22]. Once *Staphylococcus aureus* enters the bloodstream, a cascade of host events and bacterial responses can occur. First, *S. aureus* lipoproteins and others act on toll-like receptors, such as TLR-2, causing the production of pro-inflammatory cytokines TNF-α and IL-6. In addition, interactions of immune cells with peptidoglycan and lipoteichoic acids in the cell wall of *S. aureus* cause the release of IL-1 and IFN-γ [22]. The release of these factors activate complement and coagulation

pathways, initiate arachidonic acid metabolism, and cause the release of platelet-activating factor, resulting in fever, hypotension, capillary leak, and multiorgan dysfunction [23].

In order to control systemic bacteria, innate immune cells clear bacteria from the bloodstream. Therefore, *S. aureus* has developed several mechanisms in order to evade clearance. These can be grouped into two major categories: those that directly target the leukocyte, and those which indirectly target their products or mechanisms. In the first category, *S. aureus* releases a number of leukotoxins, such as the Panton-Valentine leukocidin (PVL), luekocidin AB/GH (LukAB/GH), leukocidin ED (LukED), and γ-hemolysin (Hlg), which have effects on neutrophils, mononuclear phagocytes, and T-cells [24]. Of particular note, the Panton-Valentine leukocidin subunits (LukS-PV and LukF-PV) are only detected in 5% of clinical srains, of which it is associated with 85% of community-acquired methicillin-resistant strains [24]. In the second category, *S. aureus* can evade immune cells by diminishing the C5a and N-formyl peptide receptor-driven chemoattraction via the chemotaxis inhibitory protein of Staphylococci (CHIPS) [25]. In addition, *S. aureus* can avoid phagocytosis by inhibiting C3 convertase activity via the Staphylococcal complement inhibitor (SCIN) [26, 27] and by expressing a polysaccharide capsule [22]. Coagulation is also a major factor in *S. aureus*-driven infections, and will be discussed in the following section.

Staphylococcus aureus Endocarditis: A Consequence of Invasive Infections

Among patients with *S. aureus* bacteremia in a cohort study in 1994, infective endocarditis (IE) accounted for 37% of complicated infections [28]. This condition is associated with mortality rates of 25-40% [21]. Early, sensitive diagnosis is paramount to the survival of patients in this infection. Systemic embolization occurs in 22 – 50% of cases of IE, occurring most often within the first 2 to 4 weeks of antibiotic therapy, and can affect the lungs, coronary arteries, spleen,

bowel, and extremities [21]. The occurrence of these embolisms is greatly dependent on the size of the vegetation [21].

Many virulence factors have been associated with high propensity for endocarditis. One study has indicated the methicillin-resistant *S. aureus* strain MW2 endocarditis was highly dependent on superantigen levels [29], which stimulate non-specific T-cell proliferation and activation. Moreover, our laboratory has previously described the dependence of coagulation on the formation of endocardial vegetations [30]. This coagulation is driven by two major bacterial virulence factors: staphylocoagulase (SC), and von Willebrand factor-binding protein (VWbp). Staphylocoagulase binds to the exosite of prothrombin, initiating a conformational change which exposes the fibrinogen-cleaving active site [31-33]. Similarly, VWbp has been shown to bind and activate prothrombin, although this mechanism has been shown to be through a slow conformational change that is substrate-dependent [34]. Once activated, the SC-Prothrombin initiates the cleavage of fibrinogen into fibrin while simultaneously binding to the growing vegetation on the injured heart valve [30]. Staphylocoagulase deficiency was also determined to increase the survival percentage of mice [30].

Early characterization of staphylocoagulase, as well as its molecular target prothrombin, were investigated by Dr. Morris Tager at New Haven Hospital. For these studies, they described a cutaneous wound isolate titled #104 [35-37], which we now term Tager 104. This strain has been shown to have the greatest formation of vegetations *in vivo* [30], hypothetically driven by its previously-described propensity for coagulation by Tager and others. Based on its early isolation in the 1940's, as well as its strong association with invasive disease complications such as endocarditis, Tager 104 will be elucidated as a reference strain for the systems biology study of invasive *S. aureus* infections (Chapter 4).

Staphylococcal Genomics

In line with the advent of next generation sequencing (NGS) approaches, investigators have focused on the whole genome content of *S. aureus* and other bacterial pathogens. More specifically, research has focused on characterizing the core genome, the portion of the genome containing genes vital to cell survival, and the accessory genome, the portion which encodes proteins advantageous for a particular environment [38]. In the case of *S. aureus*, most regions are highly conserved between strains, but several regions display high variability, causing *S. aureus* genomes to reach sequence variability of up to 20% [39]. This variability comes from the presence of mobile genetic elements (MGEs), which move horizontally into the strain and are then propagated vertically into subsequent generations. These MGEs can be transferred horizontally from other prokaryotes, such as pathogenicity islands, or from bacteriophages, such as prophages. One example of the latter is the clinical isolate *S. aureus* strain Newman, isolated in 1957, which contains virulence determinants which were acquired almost exclusively from phage transfer [39]. Despite this important distinctive difference in the Newman genome, it has been used extensively as an isogenic strain for the study of *S. aureus* infections. Furthermore, the modern clinical isolate *S. aureus* TW20 shows a complex amount of MGE's encoding virulence factors and antibiotic resistance cassettes [40]. Contrastingly, the *S. aureus* genome MSHR1132 shows a limited MGE content, as well as a lack of staphyloxanthin, the golden carotenoid pigment for which *S. aureus* is named, indicating its early-branching lineage [41]. Although this would seem ideal for the study of *S. aureus* infectivity, MSHR1132 is limited as an isogenic clinical strain in two ways. First, it is 10% divergent from other S. aureus strains, which themselves show a 2% or less genomic divergence [41]. Second, it is inferior in its ability to cause systemic infections, most likely due to the lack of certain MGE-encoded virulence factors required for this mode of infectivity [22, 41]. Therefore, in order to bridge the gap between strains which contain a limited MGE profile which

cannot cause systemic infections, and strains with complicated MGE profiles and strong propensity for systemic infections, research requires a strain which contains the limited set of MGE's needed for the formation of invasive infections which is: 1) 2% or less divergent from modern clinical reference strains; 2) free from the presence of antibiotic resistance cassettes; and 3) able to invade host tissue and form infections such as endocarditis. By identifying this isogenic reference strain, our results would put into context the various results of *S. aureus* whole genome characterizations (see table 4.1 for a list of current genomic strains and references) by providing this "missing link." In Chapter 4, we will investigate one potential strain, which we term Tager 104, which was isolated from an impetigo infection in 1947.

Systems Biology Approaches to Characterizing Invasive Diseases

Typically, determination of the molecular determinants of infection focus on one aspect at a time, either *in vitro* or *in vivo*. While this is initially productive in the research setting, infections in the clinical setting are multi-faceted and highly redundant. That is, infections that have many molecular determinants may either act in tandem or in redundance to achieve the goal of establishment and/or maintenance of infections. Therefore, the most informative techniques are those which incorporate all of these redundant factors in real time *in vivo* models to determine their overall contribution to pathology. Two of these holistic technologies, molecular imaging and next generation sequencing, will be utilized in this dissertation. Although seemingly two distinct technologies, they will be used in tandem in this study, with the readout from one provided as preliminary information for the alternative technology (Figure 1.2).

Molecular imaging, as defined by the Society of Nuclear Medicine and Molecular Imaging (www.interactive.snm.org), is the "visualization, characterization, and measurement of biological processes at the molecular and cellular levels in humans and other living systems." Since its

refinement, molecular imaging has provided vast amounts of information to both the clinical and preclinical settings. Due to its ease of use and access to imaging systems, optical imaging will be the focus of this dissertation. More specifically to infectious disease, optical imaging has been used in a wide range of applications. For example, fluorescence imaging of Gram positive infections in murine models have been achieved using a fluorescent analogue of vancomycin, termed vanco-800CW [42]. However, many fluorescent probes are associated with the disadvantage of decreased signal-to-noise ratios, due to tissue absorption and auto-fluorescence.

To combat this shortcoming, bioluminescence imaging (BLI) has also been explored as an inexpensive alternative in optical imaging. Luminescent enzymes, termed luciferases, have been taken from many different environmental sources, such as click beetles, fireflies, jellyfish, and the bacterium *Photorhabdus luminescens* [43]. As the host cells have no background luminescence, BLI can result in sensitive detection of bacterial burden. In addition, colocalization of BLI with fluorescent dyes can give indications of host response, such as the Cy5.5-conjugated cleavable matrix metalloprotease peptide [44], indicating inflammation. Bioluminescent imaging has been previously utilized in *S. pneumoniae* to determine the course of antibiotic therapy [45], in *S. pyogenes* to determine the colonization of nasal-associated lymphoid tissue (NALT) [46], and in *S. aureus* to detect endocarditis [30], osteomyelitis [47], and cutaneous wound involvement [48]. Therefore, optical imaging, as well as other modalities of molecular imaging, can provide information as to the timing and location of infections in real-time during *in vivo* studies.

In line with this holistic approach to infectious disease, next-generation sequencing (NGS) provides a snapshot of the entire genomic content of the bacterial cell. Rapid advancements in NGS technology allow libraries to be created, sequenced, and constructed *in situ* cheaply and efficiently. The major issue in constructing genomes is errors introduced by repeat elements,

which can cause incorrect rearrangements or collapsed regions [49]. In order to combat this shortcoming, next-generation technologies, such as Illumina MiSeq and HiSeq (http://www.illumina.com/), have designed paired-end reads, in which fragments are read forward and backward, with a known distance between the pairs provided to the algorithm [49]. In a separate approach, companies such as Pacific Biosciences (http://www.pacificbiosciences.com) have sought to overcome this shortcoming using longer read technologies, which bridge the repeat regions by scaffolding over the regions [50, 51]. In Chapter 4, we will discuss the combination of these technologies, as well as a new technology, Lucigen's NxSeq Mate Pair Library, in automating the completion of microbial genomes.

Once the genome is completed, it can provide new biology in many different ways. First, microbes can be exposed to certain environmental conditions *in vitro* or *in vivo*, and total RNA content can be isolated. This RNA content can then be sequenced using RNAseq analysis in order to determine the differential regulation of responses, such as virulence factors [52]. In addition, small nucleotide polymorphisms (SNPs) can be determined and submitted to genome-wide association studies to determine those mutations which significantly contribute to an observed phenotype. For example, vancomycin-intermediate *Staphylococcus aureus* (VISA) is a phenotype known to result from elements such as the thickening of the cell wall, as a result of mutations to multiple genes [53]. In one genome-wide association study, phenotypic data from broth microdilution and Etest (such as that in Figure 4.10) was correlated with SNPs for each of 75 clinical *S. aureus* strains to determine those which are significantly relevant [54].

Within, we will combine these two technologies in order to reveal new information about invasive Gram positive infections. Overall, the goal of these studies will be to reduce the "noise" of the infectious system, and identify the agents which are critically necessary for infection models *in*

vivo. The individual goals of each section is provided in Table 1.1. Future study can then validate these elements using traditional molecular biology and biochemistry approaches for individual factors.

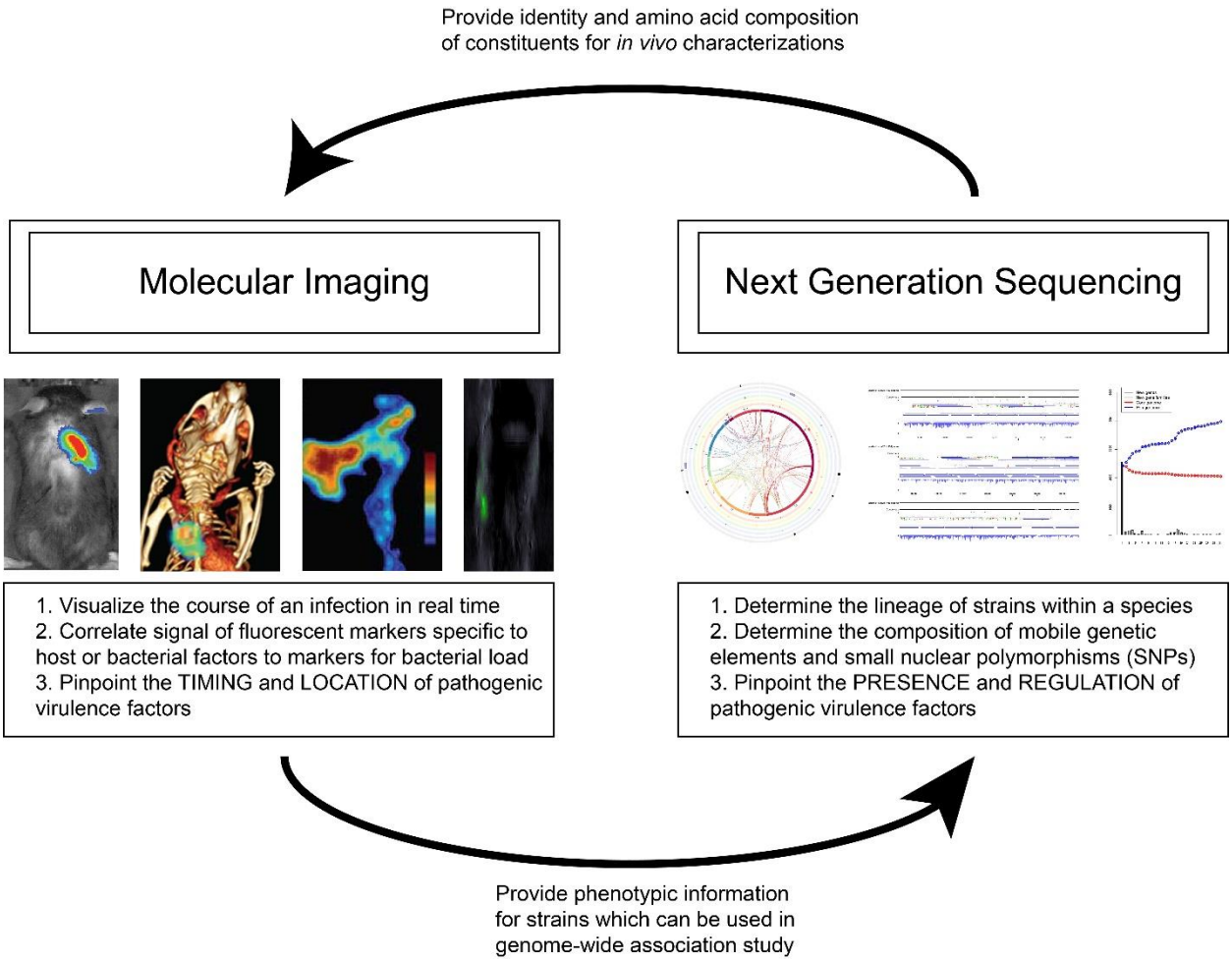

Figure 1.2: Cyclical Systeomic Approaches Using Next-Generation Technologies

Molecular imaging and next generation sequencing can work in synergy to describe the systems biology of infectious disease. First, next generation sequencing provides the identities of virulence factors contained by an organism. In turn, molecular imaging data provides information on the pathogenesis of these factors. This phenotypic information can then be submitted as quantitative data to genome-wide association studies, which determine the significance of such elements as small nuclear polymorphisms (SNPs) and/or mobile genetic elements in the establishment of the observed phenotype.

Table 1.1: Aims, Hypothesis, and Outline of This Dissertation

Associated Chapters	Hypothesis	Impact
Chapter 2: Streptokinase-Deficient Dissemination of *Streptococcus pyogenes in vivo*	*S. pyogenes* can disseminate without the use of streptokinase	Dissemination is a major event in the establishment of invasive infections and leads to further complications and death
Chapter 3: Dependence of Bioluminescence Expression on Central Metabolites in *S. pyogenes*	Repression of *S. pyogenes* bioluminescent signal is due to its dependence on catabolite repression of central metabolism in the cell	1. The ability to monitor the physiology of bacterial cells in vivo can provide useful information to future studies into antibiotic mechanisms and intracellular survival non-invasively 2. Lux-driven luminescence is a commercially available system which may be resulting in false negative results
Chapter 4: Genomic and Pathogenic Characterization of an Ancestral Staphylococcus aureus Strain Potent in Systemic Models of Infection	Tager 104, due to its early isolation and potency in systemic infections, is a potential "missing link" reference strain for systemic characterization	Identification of an isogenic, invasive *S. aureus* strain will allow the numerous completed S. aureus genomes to be placed in the proper context when studying elements of invasive disease and antibiotic resistance.

2. Streptokinase-Deficient Dissemination of *Streptococcus pyogenes in vivo*

Introduction

Streptococcus pyogenes (*S. pyogenes*), a member of the Group A Streptococci (GAS), is an aggressive human-specific pathogen that causes infections that range in severity from skin infections, such as pyoderma, impetigo, and cellulitis, to toxic shock syndrome, glomerular nephritis, and necrotizing fasciitis [3]. The global impact of GAS infections is estimated to affect 18 million people, resulting in 517,000 deaths per year [55]. *S. pyogenes* expresses proteins that are covalently tethered to the cell wall for the purpose of binding host factors, including the fibrinolytic zymogen plasminogen (Pg) and the major blood clotting protein fibrin(ogen). In normal physiology, Pg is activated by proteolytic cleavage to plasmin (Pm) to degrade fibrin clots. In the context of *S. pyogenes* infection, the pathogen secretes a bacterial Pg-activator called streptokinase (SK). GAS pathogenicity is now thought to be dependent on SK binding to Pg and inducing formation of an active SK•Pg* catalytic complex by the molecular sexuality mechanism, which then recognizes a second Pg molecule as a substrate and converts it proteolytically into plasmin (Pm) in the vicinity of pathogen [56-60]. Different *S. pyogenes* strains express convergent types of receptors to bind and employ the catalytic SK•Pg*/Pm complexes, essentially coating the bacterial surface, promoting dissemination through protective fibrin barriers produced by the initial host response to infection, and indirectly activation of matrix metalloproteinases [18]. Most available data regarding the underlying mechanism of Pg activation by SK are derived from *in vitro* studies. However, with the recent advances in non-invasive mouse imaging, it is now possible to serially track a developing infection as a function of time, allowing us to determine the *in vivo* relevance of previously reported in vitro data [2]. Here, our goal was to follow streptococcal infection in vivo, using side-by-side comparison of multiple strains to support future therapeutic

testing and imaging agent design. In addition, we test the dogma that mice are resistant to streptococcal infection due to a lack of pathogen-mediated Pg activation by SK.

Bioluminescence-expressing bacteria can be used to track populations of a pathogen *in vivo* through light production inside the cells of these microbes, driven by over-expression of either the *Photorhabdus luminescens lux* operon (*luxABCDE*) or the firefly luciferase (*ffluc*) plasmid. The most widely used instance of *lux* operon incorporation is through initial transposon insertion, followed by positive and subsequent negative antibiotic selection protocols and a temperature-shift to promote loss of the unincorporated carrier plasmid [45]. This labor-intensive process selects for bioluminescent pathogens that have all the attributes of the parent strain, aside from the damage generated at the transposon insertion site and any shifts in the open reading frames that this insertion causes. More recently, solely plasmid-driven bioluminescent *S. pyogenes* strains have been developed through use of the ω-ε-ζ operon or toxin-antitoxin (TA) system [61]. Here, the loss of the plasmid results in post-segregational killing of the bacterium during replication, such that bacteria that have the toxin but no antitoxin will ultimately result in cell death. The resulting TA-based *ffluc* plasmid has an advantage in that it is significantly smaller than the ~18 kb pXen05 carrier of the *luxABCDE* operon, and, as expected, the smaller plasmid exhibits higher transformation efficiency. In either case, maximal light producers for *luxABCDE* or *ffluc* strains are selected and tested for microbial fitness in comparison to parental strains, thereby eliminating unwanted disruption of housekeeping gene networks.

Bioluminescence as a proxy for bacterial cell burden has been reported for *Staphylococcus aureus* in models of endocarditis [62], osteomyelitis [47], and cutaneous wound involvement [48]. In regard to other Gram-positive pathogens, like *S. pyogenes*, the literature on the use of bioluminescence to track the pathogen is not as expansive, possibly because of the inherent

difficulty in genetically modifying *S. pyogenes*, given the size of aforementioned *lux* operon-carrying pXen5 plasmid [63]. For those streptococcal strains where genetic modification has been achieved, light production can also be hampered by the relatively high GC contents of the genes encoded during this process and/or the overall strength of the ribosomal binding sites [63]. Lastly, the detection limit of these engineered bacteria *in vivo* is limited by the availability of cofactors necessary for the oxidation reaction, such as the flavin mononuclease ($FMNH_2$), ATP, and molecular oxygen. Taken together, there are many factors that may mitigate the successful production of light *in vivo,* rendering the development of this promising technology non-trivial. Some examples of bioluminescent *S. pyogenes* imaging have involved colonization of the nasopharyngeal tract and began with the detection of a mucosal serotype M49 strain, termed strain 591, in surrounding nasal-associated lymphoid tissue [46]. This led to the attempt to monitor other isolates, such as the clinical *emm75* isolate, to determine the efficacy of vaccines in the murine nasopharynx [64]. The primary reason for the limited number of isolates and studies may lie in the inherent difficulty in genetically modifying *S. pyogenes*, given the size of aforementioned *lux* operon-carrying pXen05 plasmid [63]. For those streptococcal strains where genetic modification has been achieved, light production can also be hampered by the relatively high GC contents of the genes encoded during this process and/or the overall strength of the ribosomal binding sites [63]. In addition, previous studies have determined plasmid expression of the *lux* operon to be higher than that of genomic incorporations (26 photons s^{-1} CFU^{-1} versus 1.0 photons s^{-1} CFU^{-1}, respectively), although passage *in vitro* without antibiotic challenge decreased the numbers expressing this plasmid one thousand-fold [64], emphasizing a trade-off between genetic stability and luminescence expression. Lastly, the detection limit of these engineered bacteria *in vivo* is also limited by the availability of cofactors necessary for the oxidation reaction, such as the flavin

mononuclease ($FMNH_2$), ATP, and molecular oxygen. Taken together, there are many factors that mitigate the successful production of light *in vivo*, and development of this technology is not trivial.

Compounding these issues is the ample evidence that SK secreted by *S. pyogenes* lacks the ability to activate mouse Pg. Presumably to cause the spread of *Streptococcus in vivo* in humans, the pathogen secretes the Pg activator SK, allowing formation of active SK•Pg/Pm complexes that bind surface receptors, namely glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [65, 66], Pg-binding group A streptococcal M-like protein (PAM) [67-69], enolase [70, 71], and other bacterial surface proteins [72-74]. The nature of the SK•Pg*/Pm receptor is dictated by the *S. pyogenes* strain. Because of this, the molecular mechanism driving Pg activation by SK and their respective interaction with the bacterial receptors has been vigorously studied [18, 19, 73, 75-81]. The rationale for focusing on SK relies partly on the observation that streptococcal infections have increased mortality in transgenic mice that express 14-16% plasma levels of the human Pg [82] and in mice given intraperitoneal human Pg bolus injections [18]. Despite these studies, it remains unclear whether Pg activation is a strict requirement, and whether lack of this activation potential precludes disseminative infections. This lack of mouse Pg activation may impart some level of resistance to mice against *S. pyogenes* infection.

Here, we utilized wild type mice (i.e. C57BL/6) that should be devoid of the enhanced *S. pyogenes* lethality seen in the humanized Pg transgenic mice. We tracked *in vivo* dissemination though non-invasive longitudinal imaging studies by use of pathogens containing the aforementioned *lux* operon. For this study, we compared the disease progression after infection with popular streptococcal strains using imaging and histological analysis. *S. pneumoniae* Xen10, derived from the parent strain A66.1 Serotype M3, has previously been studied in pneumonia, sepsis, pouch,

and meningitis models [45, 83-94]. For the bioluminescent *S. pyogenes*, we used the aforementioned Xen20 that was derived from the parent 591-serotype M49 strain in nasopharyngeal or impetigo murine models [46, 95-97]. We also utilized the classical *S. pyogenes* ALAB49 serotype M53 strain (*emm* pattern D or a "skin specialist" strain) [98], as a non-luminescent reference. Of note, both ALAB49 and the Xen20 (*emm* pattern E or designated a "generalist" strain that denotes equal parts "skin specialist" and "throat specialist" strains) were isolated from serious human infections, namely impetigo (Alabama, 1986) and acute post-streptococcal glomerulonephritis, respectively [99, 100]. Our side-by-side comparison demonstrated that despite the lack of SK-associated Pm production by the pathogen, *S. pyogenes* disseminates from an initial subcutaneous infection, leading to systemic disease and multiple organ involvement for both ALAB49 and Xen20, as determined by post-mortem assessment of the organs for colony forming units (CFU) determination and Gram-staining. Interestingly, bioluminescence imaging did not recapitulate these findings, as the Xen20 showed no signal outside the initial site of infection despite exhibiting higher mortality than the *S. pneumoniae*-infected mice. Moreover, our studies demonstrate incorporation of the *lux* operon was done under a glucose-sensitive promoter, which may explain our divergent results as the pathogen seeds different organs. Taken together, our results indicate a Pg-independent pathway for *S. pyogenes* dissemination. Further, we conclude that *in vivo* characterization of *lux* operon insertion into *lux* operon-driven bioluminescent pathogens is necessary prior to their use as reliable proxy for bacterial burden assessment.

Materials and Methods

Animals

For this study, 36 C57BL/6 mice were purchased from The Jackson Laboratories (Bar Harbor,

ME). Mice were housed at the Auburn University College of Veterinary Medicine with *ad libitum* access to alfalfa-free chow and water. All procedures were designed in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and approved by the Institutional Animal Care and Use Committee of Auburn University.

Plasminogen Activation Assay

The protocol used was described previously [39]. Briefly, the supernatants of cultures of Xen20, Xen10, and ALAB49 at mid-log phase (λ_{600nm} of ~ 0.6) were harvested, filtered, and frozen at -80 º C until use. The aliquots were thawed, and the following components combined for each assay (all concentrations noted are final): human Glu-Pg and fibrinogen (Fbg) when added at concentrations of 20 µg/mL, 40 µL of supernatant per strain, substrate S-2251 (*D*-Val-Leu-Lys-para-notroanilide; Chromogenix, Thermo Fisher Scientific, Pittsburgh, PA) at a concentration of 0.17 mM, with the addition of 50 mM tris(hydroxymethyl) aminomethane pH 7.5 (Tris buffer) to a final volume of 300 µL. For the positive control, human Glu-Pg at a concentration of 20 µg/mL, purified SK at a concentration of 118 nM, and substrate S-2251 at a concentration of 0.17 mM were combined with the addition of the Tris buffer until a final volume of 300 µL was achieved. For the activity assays containing human Fbg, the human Glu-Pg and Fbg were pre-incubated in the Tris buffer for 25 minutes, before the addition of supernatant for the remaining 15-minute incubation. The samples were read by a SpectraMax® 340PC^{384} (Molecular Devices, Sunnyvale, CA) at an absorbance of 405 nm every 37 seconds for 1 hour.

Preparation of Bacterial Injections

Strain Xen20 and ALAB49 was grown in Brain Heart infusion (BHI) broth (Research Products International (RPI), Mt. Prospect, Illinois) and Strain Xen10 in BHI broth supplemented with 5% sheep blood (Northeast Lab Services, Waterville, ME), both with addition of 200 µg/mL of

kanamycin, for 18 hours at 37 °C. Cells from stationary cultures of 1½-2 day cultures were pelleted, washed in phosphate buffered saline (PBS; Boston BioProducts, Ashland, MA) with 10% glycerol twice, and resuspended in PBS with 10% glycerol for injection with the addition of 1 mg/mL of sterile Cytodex beads (Sigma-Aldrich, St. Louis, MO).

Animal Procedures

To study the extent of bacterial dissemination throughout the mice, we induced infections in C57BL/6 mice by subcutaneous injection of between 3 and 5 x 10^7 colony forming units (CFU) of streptococcal bacteria 24 hours prior to the first imaging session and imaged daily from days 2-6, as detailed in Figure 2.1a. Twelve mice were used for each strain (Xen20, Xen10, and ALAB49), for the 36 aforementioned mice. For these studies, mice were anesthetized by isoflurane (Piramal Healthcare, Boston, MA) with 1-3 % medical grade oxygen using a XGI-8 Anesthesia System (PerkinElmer, Waltham, MA) and the hair between the shoulder blades was removed by acute treatment with Nair (Church & Dwight Co., Inc., Ewing, NJ). The mice were imaged every 24 hours for 5 minutes with the IVIS Lumina XRMS Imager (PerkinElmer) using a binning of four and default aperture settings. The mice were then sacrificed on day 8 post infection and the heart, lungs, liver, kidneys, spleen, and dermis area surrounding the site of injection were collected, subjected to bioluminescent imaging as described above, and subsequently fixed in 4% paraformaldehyde in PBS prior to histological analysis or weighed and placed in sterile PBS for homogenizing.

Histology and Ex Vivo Imaging

Tissue was rinsed with sterile PBS and dabbed dry on a gauze pad prior to positioning in partially filled disposable embedding forms. The remainder of the exposed tissue was immersed in VWR Clear Frozen Section Compound prior to submersion and freezing over an isopentane (Alfa Aesar,

Ward Hill, MA) bath maintained over dry ice. The block was maintained for long-periods at -40 °C and equilibrated to -20 °C prior to sectioning using a Cryostat (Thermo Scientific, Kalamazoo, MI). Slices were typically between 8-12 microns in thickness, depending on tissue hardness. Alternating sections were stained with either H&E or Gram stain by standard protocol and the cover slip attached by applying 2-3 drops of Cytoseal (Thermo Scientific). Histology images were captured with on an Axioskop 50 (Carl Zeiss Microscopy, Jena, Germany) with a 0.76x, C-Mount adapter and using a DS-Fi1 camera head (Nikon, Melville, NY) with DS-L3 control unit. The tissue collected for homogenizing was weighed and then placed into 5 mL of either sterile PBS or 0.85% saline solution and homogenized with a Tissue-Tearor (BioSpec Products, Inc., Bartlesville, OK). The resulting solution was serially diluted to concentrations ranging from 10^{1}-10^{-12} and plated on tryptic soy agar (TSA) supplemented with 5% sheep blood for the strain ALAB49 or BHI agar supplemented with 5% sheep blood and 200 μg/mL of kanamycin for the strains Xen20 and Xen10. After 18-hour incubation, the plates were imaged with a 1-second integration time with the IVIS Lumina XRMS Imager.

Glucose Dependence Assay

For assessment of the relative effect that *D*-glucose has on the bioluminescence production, 1 mL of Xen20 cells were grown to stationary phase in BHI broth containing 200 μg/mL kanamycin, confirmed by an optical density measurement ($OD_{600\ nm}$) $\geq$ 1, which did not fluctuate. Cell supernatants were removed after centrifugation and the cell pellet re-suspended in M9 minimal medium [101] without *D*-glucose. To achieve the range of *D*-glucose concentrations (0-6 mM), cells were combined with *D*-glucose to final concentrations between 0-6 mM. Samples were placed at an inclined position within the IVIS Lumina XRMS on the 37 °C stage and imaged with a kinetic protocol. The imager was set to automatically adjust exposure times at 20-minute time

points, for a total of 8 time points spanning 140 minutes. The flux (photons/second) of each sample was obtained from region of interest (ROI) for each sample using Living Image software (PerkinElmer). Statistical analysis was performed using Repeated Measures Two-Way Analysis of Variance (ANOVA) on R-statistical software version 3.1.2 (http://www.r-statistics.com/on/r/).

Results

Activation of Human Plasminogen by S. pyogenes Xen20 and ALAB49 Supernatants

Activation of human Pg by streptococci is a virulence determinant and involved in dissemination of the bacteria through host tissue [82]. While activation of human Pg has been well studied, we here characterized the Pm generation potential and underlying mechanisms of human Pg activation for given strains while comparing to positive and negative controls. Figure 2.1 demonstrates the relative activation of human Pg by SK from *S. pyogenes* ALAB49 and Xen20 in the presence of human Fbg. Xen10, which is a SK-deficient strain. Purified SK (118 nM) were used as negative and positive controls, respectively (Figure 2.1A). The maximum absorbance reached was observed as an asymptote at 1.0, reached at time point 30 minutes, for both Xen20 and ALAB49 strains, as shown in Figure 2.1B and Figure 2.1C, respectively. Therefore, SK-containing supernatants from ALAB49 and Xen20 show minor differences in their ability to activate Pg in the absence of Fbg (i.e. ~7% for Xen20) but similar enhanced activation of human Pg in the presence of Fbg. Of note, the identical experiments using mouse Pg instead of its human counterpart showed no activation of murine Pg regardless of Fbg supplementation (data not shown).

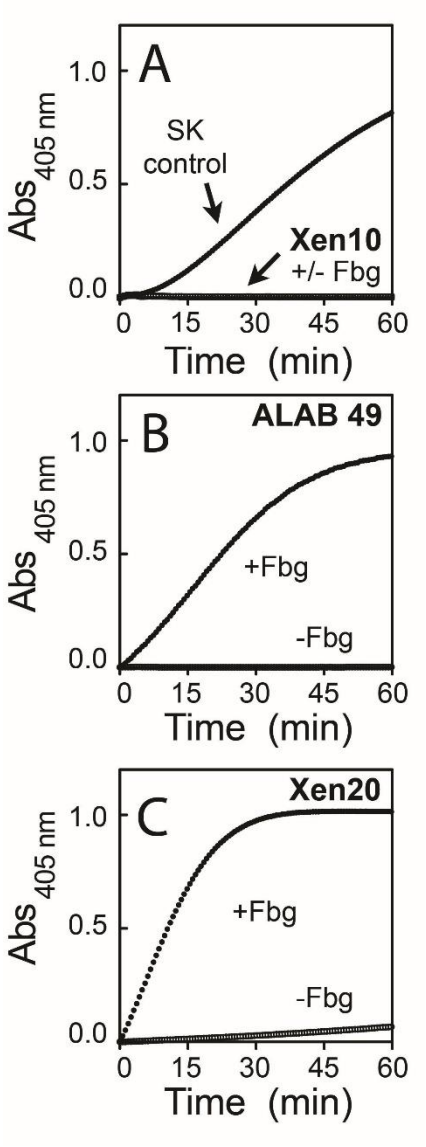

Figure 2.1: Human Plasminogen Activation by the Streptococcal Secretomes.

A, Activation of human plasminogen (Pg) with and without fibrinogen (Fbg) for supernatants of log-phase streptokinase (SK)-deficient *S. pneumoniae* (Xen10) and a SK (118 nM) positive control. *B* and *C,* Activation of Pg with and without Fbg for supernatants of *S. pyogenes* strains ALAB49 and Xen20, respectively.

Genomic Analysis of S. pyogenes Plasminogen Activation System

The dissemination of *Streptococcus pyogenes* in humans and human Pg-expressing mice is dependent on the activation of Pg by SK. Since there is some overlap in the phenotype of *emm* pattern D and E streptococcal strains, we first characterized the Pg activation system in both ALAB49 and Xen20 prior to the *in vivo* mouse challenge that would abrogate these differences. To our knowledge no other study has characterized the exact SK allele expressed by Xen20. Phylogenetic analyses of SK genes demonstrated three clusters 1, 2a, and 2b [102]. SK of cluster 1 (SK1) demonstrates the ability to activate soluble Pg without the obligatory need for Fbg supplementation to form the SK1•Pg*•Fbg•M1-protein complex on the pathogen surface. In contrast, SK of cluster 2 (SK2a or SK2b) can only activate Pg pre-bound to Fbg through the streptococcal M1 coat protein, allowing formation of the SK2b•PAM-protein complex and providing Pg to the SK2b as a neighboring (Pg/Pm)•Fbg•M1-protein complex formation. ALAB49 belongs to SK2b and the results of our kinetics mirror those reports (Figure 2.1). The equivalent response of both ALAB49 and Xen20 supernatants to Fbg supplementation suggests that Xen20 is the SK2a allele evident by some activation of human Pg (~7%) in the absence of Fbg. Others have noted that there is a strong inheritance pattern between PAM and SK2b in strains specific for skin infections such as ALAB49 [103]. Therefore, we investigated whether the parent strain of Xen20 denoted as a strain 591, serotype M49 strain has a gene homologous to PAM through BLAST alignment with the PAM protein sequence from ALAB49 (Acc. No. **G4R2S6)**. Although no homology between the sequences was identified, the existence of another cell wall Pg binding protein was identified. The extracellular protein factor (Epf) (Acc. No. **B5XKN1**), originally identified in serotype M49 and confirmed in strain 591, has been shown to bind immobilized Pg [72]. The presence of Epf has been shown to increase the ability of strain 591 cells

to adhere to and internalize into human keratinocytes and binds Pg with a determined K_d value of 0.28 μM, which is physiologically relevant, as Pg circulates at a concentration of 2 μM [73]. Our results indicate that Xen20 and ALAB49 express cluster 2-type SK, yet Xen20 is an SK2a and ALAB49 is an SK2b. Therefore, *in vivo* comparison of their relative infectivity may eventually discriminate these cluster types through molecular imaging efforts.

Diminished Bioluminescence at the Site of Infection for Xen10 and Xen20 strains

To study dissemination of these *Streptococcus* species in wild type mice, we monitored bioluminescence production to determine whether the absence of productive Pg binding and activation would limit spread of the pathogen to distal organs. To accomplish this, mice were injected subcutaneously with bacteria and imaged daily, as illustrated in Figure 2.2A. A representative imaging signal time course is shown in Figure 2.2B. Despite the decrease in bioluminescence observed for *S. pyogenes* Xen20, these infections resulted in 33% mortality (4 of 12 mice). In contrast, all mice survived *S. pneumoniae* Xen10 infections (0 deaths in 12 infected mice) and ALAB49 infections had the highest mortality rate of 50% (6 of 12 mice) in the span of 8 days. By bioluminescence imaging, the *S. pneumoniae* infections in the different biological replicates appeared resolved at day 5 (Figure 2.2B). The bioluminescence associated with Xen20 infection site diminished over time but did not show dissemination of the signal or other organ involvement. The continued presence of the bioluminescent signal from these pathogens over days led us to question the state of other organ involvement via histology of organs recovered from necropsy.

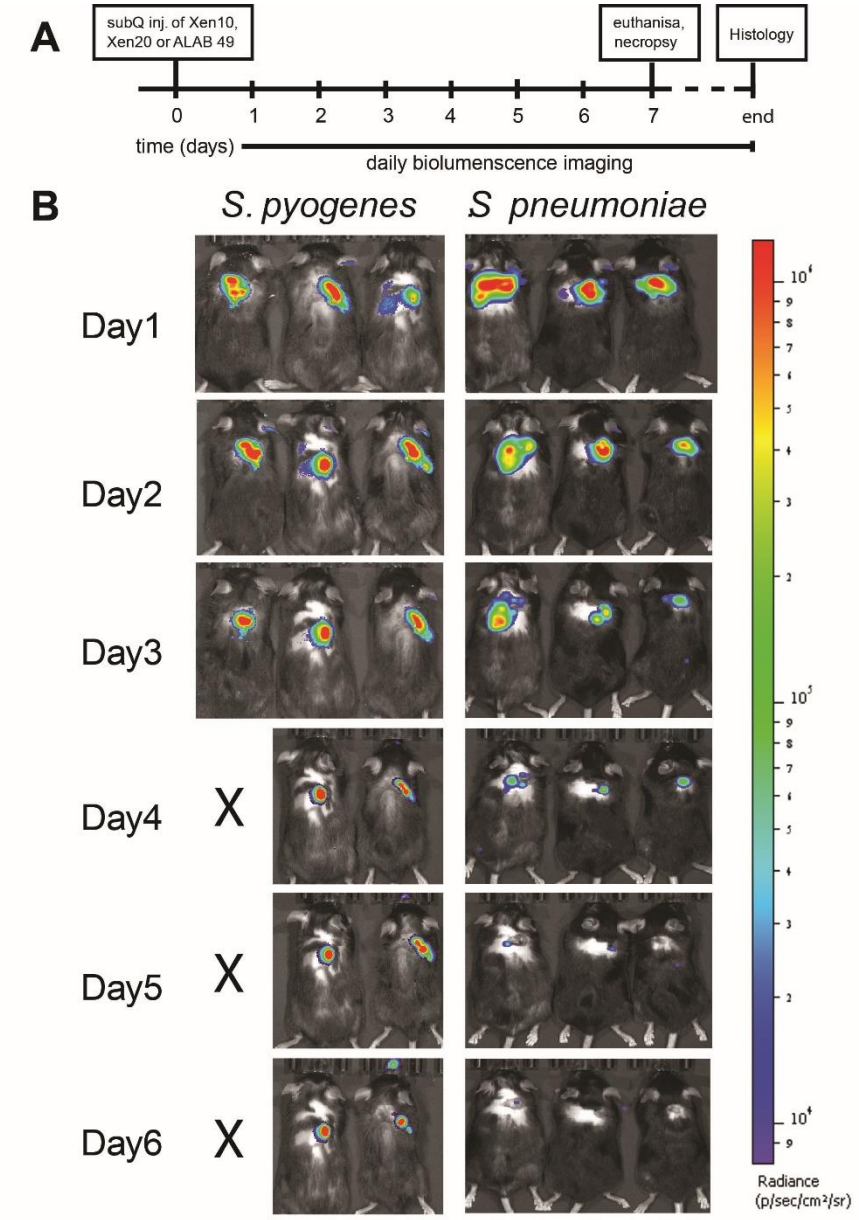

Figure 2.2: Time course of Streptococcal Infection Monitored by Bioluminescence Imaging.

A. Experimental Design. *B.* Daily imaging of mice after subcutaneous injection of *S. pyogenes* (Xen20) and *S. pneumoniae* (Xen10). Bioluminescence imaging for a representative of 3 mice for the light producing microbes, shown at the indicated time points, n = 12 for each group.

Gram Staining and CFU Determination for Pathogens in Distal Organs and the Skin

To address the distal organ involvement in these streptococcal infected mice, we performed histology and analyzed Gram-stained slides from each cohort. In all mice, we detected signs that an infection had occurred (Figure 2.3). For Xen20 and ALAB49 infected mice (Figure 2.3A and 2.3B, respectively), there was apparent dissemination from the initial skin infection to several organs, including the kidneys, heart, liver, and lungs. ALAB49 had, on the average, higher relative content of the pathogen in the kidney and heart than its bioluminescent counterpart Xen20 by gross inspection (Figure 2.3B). Interestingly but not entirely unexpected, Xen10 showed the greatest bacterial burden in the lung compared to the other streptococcal strains, with lesser was pathology in the kidney, spleen and heart. Examination of the Gram stained slides for Xen10 under 400-times magnification showed Gram-positive cocci were observed as clusters or debris within splenocytes as opposed to the *S. pyogenes* infections (Fig. 2.3C). Therefore, Xen20 and ALAB49 infections have the capacity to overwhelm the host immune defenses, even in the absence of productive Pg activation.

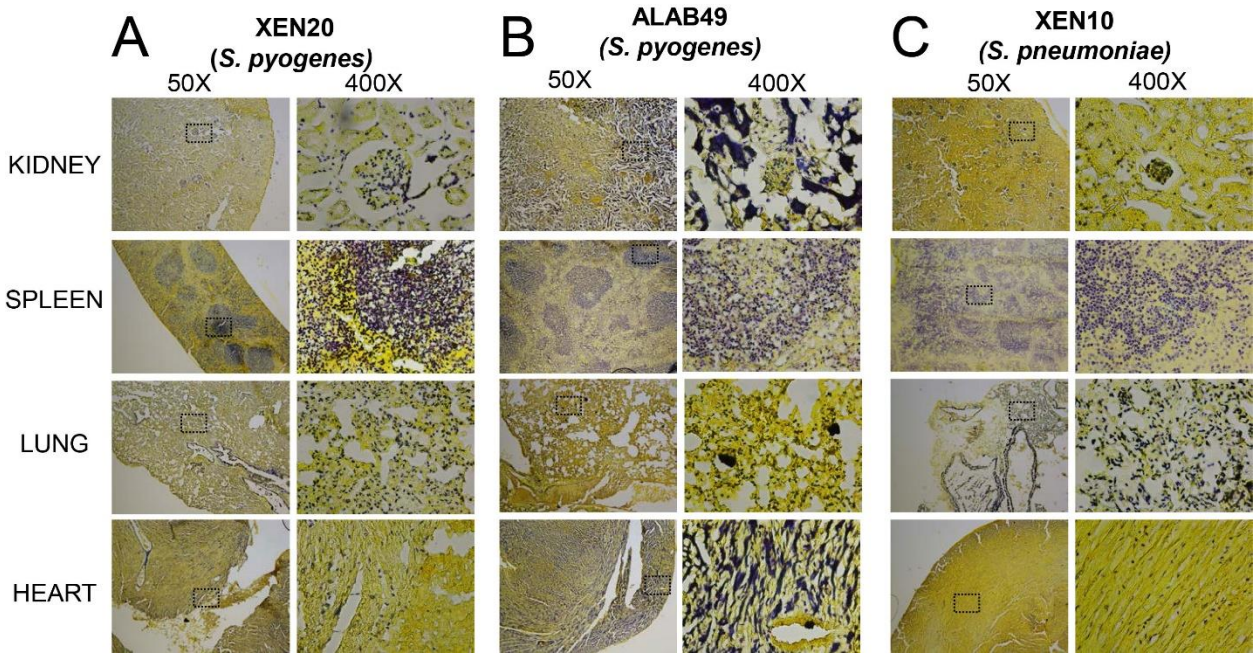

Figure 2.3: Histological analysis of Streptococcal strains reveals widespread dissemination.

Histological analysis of isolated organs Gram-positive *Streptococcus* species (*blue*) and tissue (*yellow*) are shown at 50 times magnification, with the area of interest (*black rectangle*) shown at 400 times magnification. Representative data from 2 injections of 2 mice from each group for a total n = 4.

To confirm the Gram staining, we quantified bacterial numbers or content within an organ. This classic CFU determination per gram of isolated tissue allowed us to verify metabolically active and living cells (Figure 2.4). We produced homogenates of the organs, plated dilutions, and counted the number of viable *S. pyogenes* cells (Figure 2.4A-B). Only trace colonies were detectable at the maximum Xen10 content (data not shown), but these cells were likely contaminants as we observed no concurrent bioluminescent signal in these plates. As indicated by the graphs in Figure 2.4A and B, there was a significant pathogen CFU amount, yet whole organ bioluminescence imaging did not recapitulate these results, despite the presence of 10^5-10^7 CFU per organ, even at longer exposure times of up to 10 minutes (Figure 2.4C). To eliminate destruction of the *lux* operon or contaminating bacteria as a cause of these results, we imaged bioluminescence in selected plates. Our results showed individual colonies with levels of calibrated luminescence corresponding to the peak signal for this pathogen provided by the vendor (Figure 2.4D). Therefore, we concluded that instability of the *lux* operon during the course of infection was not the cause for the loss of *in vivo* bioluminescent signal.

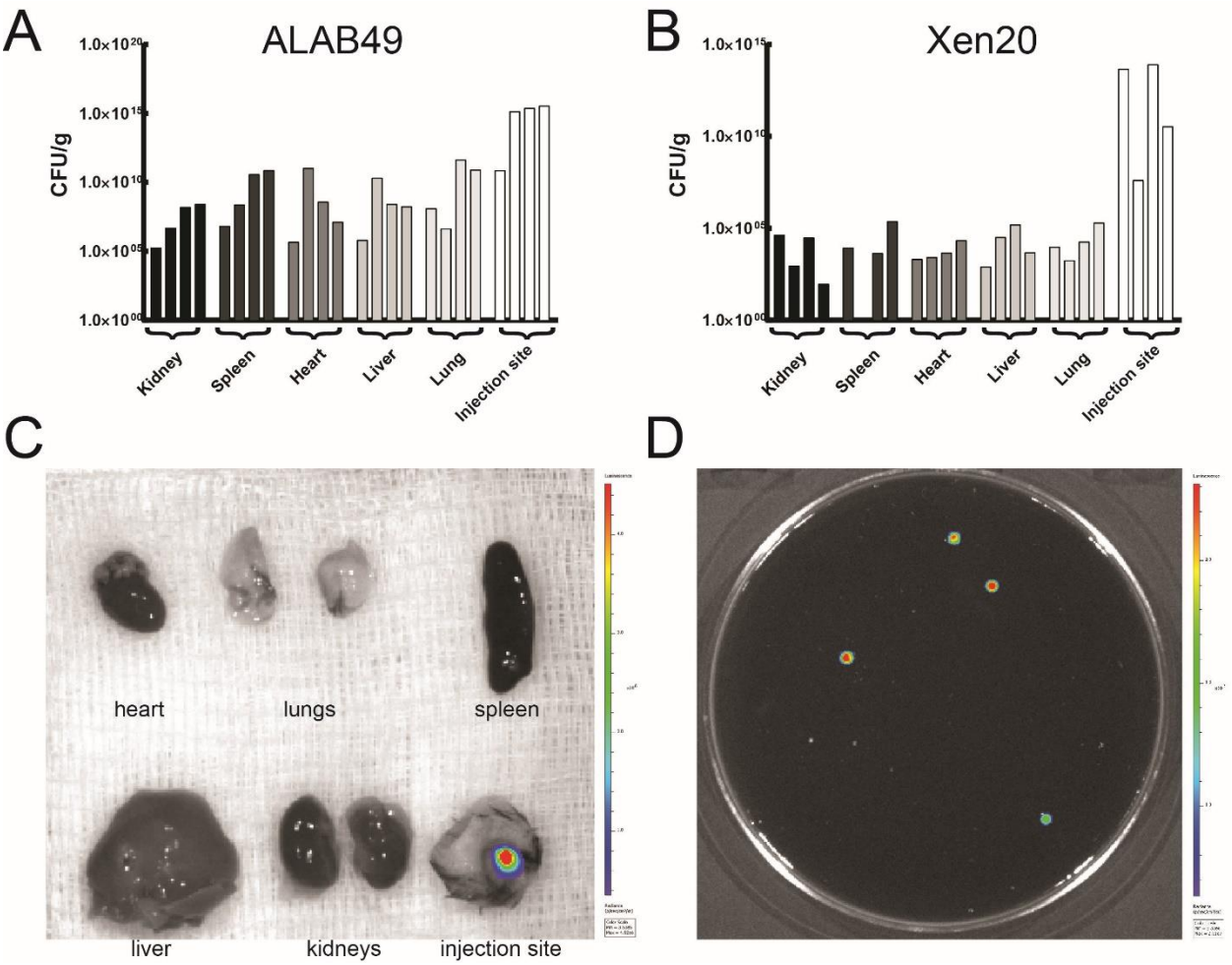

Figure 2.4: Evidence of bioluminescent colonies retrieved from *in vivo* infection.

A-B. Mean colony forming units (CFU) per organ recovered from homogenates of 2 separate injections of 2 mice each for ALAB49 (*A*) or Xen20 (*B*), with n = 4. *C.* Bioluminescence from organs (from left to right): heart; lungs; spleen; liver; kidneys; infection site. Signal is only observed from the skin at the infection site, despite CFU's ranging from $1x10^5$ – $2x10^7$ present within the organs. *D.* Bioluminescent colonies of Xen20 recovered from the spleen from Panel b.

Presence of Exogenous Glucose Modulates Xen20 Bioluminescence in a Concentration Dependent Manner

To determine if the *lux* operon was inserted into a region regulated by a particular host factor, we analyzed the genes disrupted by the transposon insertion of the bioluminescent cassette by BLAST search analysis. The plasmid pXen05, formerly termed pAULA-Tn*4001 luxABCDE* Km^R, was previously shown to have integrated at a single site in the Xen20 parent, strain 591, in the second open reading frame for a hypothetical protein (Acc. No. NP_268809) [46]. BLAST search of this hypothetical protein is 100% homologous to a protein identified in the M1 strain of GAS as a sugar phosphate isomerase or epimerase (*iolE*) involved in carbohydrate transport and metabolism. Fluorescence microscopy studies of green fluorescence protein (GFP) driven under the *iolE* promoter in *Salmonella enterica subspecies 1 serovar Typhimurium* strain 14028 revealed strict repression of GFP-expression when grown on *D*-glucose rich medium [104]. Therefore, we assayed the bioluminescence production of Xen20 cells, washed and re-suspended in M9 minimal medium with a range of *D*-glucose concentrations from 0-6 mM (Figure 2.5A). This medium limits the secondary effects of carbon sources in the medium, allowing an accurate depiction of glucose dependence. Quantification of this luminescence expression revealed the addition of low concentrations of *D*-glucose increased the luminescence observed in the tubes up to a concentration of 4.0 mM and decreased up to 6 mM (Figure 2.5B). Replicates of measurements shown in Figure 2.5B were combined and presented as the percentage of peak average flux to correct for individual replicate variance, and graphed for each time point and glucose concentration (Figure 2.5C). Repeated measure two-way ANOVA was performed using the statistical software R 3.1.2, and revealed that both time and glucose concentration, as well as the interaction of these two variables, had a significant effect on light production in Xen20 ($p > 0.05$). Luminescence is

recoverable by plating these separate reactions onto BHI plates and incubating for 16 hours, excluding loss of the *luxABCDE* cassette and no cell death during the 160-minute incubation with glucose (data not shown).

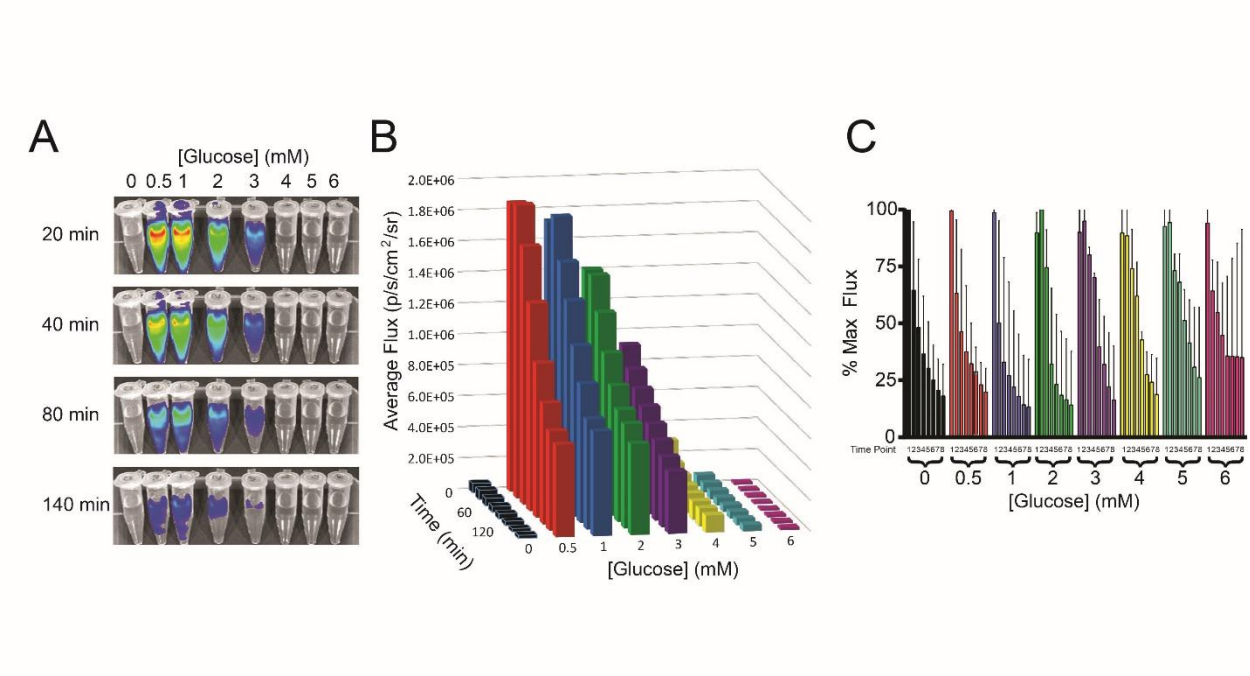

Figure 2.5: Modulation of bioluminescence as a function of glucose concentration in a light producing *S. pyogenes*.

A, Xen20 cells in stationary phase (OD_{600nm} > 1) were washed and resuspended in M9 minimal medium and combined with indicated *D*-glucose concentration. Shown is one representation of 3 trials. *B*, Quantification of luminescence in average flux of samples depicted in *a* versus time for a total of 140 minutes. *C*, Maximum average flux for each glucose concentration was divided by the average flux at each time point in order to produce the percent maximum expression for each of three trials. Data shown here are an aggregate of 3 trials.

Discussion

Our goals in this study were to determine whether there is an *in vivo* requirement for host Pg activation for the spread of *S. pyogenes* to distal organs. Further, we explored bioluminescence imaging of *S. pyogenes* Xen20 to non-invasively follow the spread of a cutaneous infection to other organs. We determined that (i) the SK from Xen20 and ALAB49 are cluster 2a and 2b alleles, respectively; (ii) that there was a distinct Pg-independent pathway that allows *S. pyogenes* to spread *in vivo* for both Xen20 and ALAB49; (iii) that there was an increased mortality in cohorts infected with the *S. pyogenes* Xen20 and ALAB49 compared to the *S. pneumoniae* Xen10; (iv) that the insertion site of the bioluminescent *lux* cassette into the Xen20 genome may be under control of a sugar phosphate isomerases/epimerases (*iolE*) promoter; and (v) that the light generation by the *S. pyogenes* Xen20 is sensitive to the local *D*-glucose concentrations leading to a masking of pathogen content in distal organs. In the discussion below, we relate each result to *in vivo* determinates of streptococcal infections to aid future non-invasive imaging studies.

To focus our studies on *S. pyogenes* dissemination *in vivo*, we asked whether longitudinal and non-invasive imaging of Xen20 correlated with disease progression and if we could use the bioluminescent signal from Xen20 as a surrogate for overall bacterial burden. Because virulence of streptococcal pathogens is related to the expression of microbial virulence factors, including exotoxins (i.e. streptolysin O and S, streptococcal pyogenic exotoxin A and C), hyaluronidase, C5a peptidase, streptococcal chemokine protease and SK, we evaluated streptococcal infections in a Pg-independent model of cutaneous infection, thereby testing whether activation of host Pg is strictly required for *S. pyogenes* virulence. One example of these virulence factors is the cysteine protease SpeB, which is found in all GAS strains and cleaves cytokine precursors, immunoglobulins, and immune system-associated receptors [75]. Inactivating this protein decreases the dissemination of GAS. Contrary to this, SpeB also inactivates SK, thereby

modulating Pg activation potential of the pathogen and limiting its spread [77, 105]. Streptococcal deoxylribonuclease, called streptodornase, and phage-acquired DNases function to clear the infection site to aid in dissemination and immune system evasion [75]. Therefore, we sought to determine if these factors, along with others, are sufficient to allow for the spread of the *S. pyogenes* infections in the absence of a productive SK-Pg system. This model is not deficient in Pg, as Pg-deficient mice experience aberrant developmental processes of unknown consequences, but rather employed a wild type mouse, where the activation of Pg by SK is not possible [106]. As a result, there is a potential for formation of an inactive mouse SK•Pg complex that may be localized to the pathogen surface through microbial plasmin(ogen) binding proteins, which should mimic the state of the microbe in the human disease progression.

We compared two *S. pyogenes* and one *S. pneumoniae* strains by assessing the migration of the light source as the disease developed. To better understand the *S. pyogenes* strains that we chose, initial characterization of ALAB49 and Xen20 secretomes was performed using both human and mouse proteins. Both ALAB49 and Xen20 belonged to the cluster 2 alleles of SK (Figure 2.1), but genomic analysis revealed that the anchoring of the Pm to the surface of the pathogen was different in that Xen20 is PAM-negative (SK2a) and ALAB49 is PAM-positive (SK2b). Since Fbg-enhanced human Pg activation by both pathogens was similar, it is reasonable to assert both pathogens would be impaired in a similar way *in vivo*, given the lack of mouse Pg activation. Because ALAB49 is a skin specialist strain and Xen20 is a generalist strain, our side-by-side comparison between these two *S. pyogenes* strains, was done at the phenotypic imaging level, asking the question whether the same inoculum lead to abscess formation in the kidney or other organs, and whether or not bioluminescence recapitulates the bacterial burden in distal organs.

Our *in vivo* results were somewhat unexpected in that the total bacterial burden in the distal organs determined by CFU counts and histology was not recapitulated by our bioluminescence imaging data (Figure 2.3). We saw no signal originating in the vicinity of the distal organs (i.e. liver, kidney and spleen) for the Xen20 pathogen-infected mice, despite a significant 50% mortality for both the *S. pyogenes* strains (Xen20-ALAB49) on day 7. Indeed, excision of these organs, in an effort to limit tissue penetration effects, revealed no bioluminescent signal. These results are similar to that determined by others, in which dissection and removal of the bladder led to no detectable levels of bioluminescent signal [63]. This lack of light production in the infected organs could be caused by several factors, including (i) reduction of cellular molecular oxygen to catalyze the redox reaction by which luminescent substrates are oxidized to produce light, (ii) diminished presence of reducing equivalents (i.e. $FMNH_2$ and /or ATP) in stationary phase growth seen by the microbe in tissues, or (iii) exhausted substrate pool for bacterial luciferase due to shifts in carbon source bioavailability. As previously mentioned, *S. pyogenes* was also developed using a TA-based *ffluc* plasmid system, and showed a greater signal than standard *lux* plasmid, a result which was previously speculated to be due to substrate availability and/or promoter regions [61]. Taken together, our data show that even in the absence of Pg activation both *S. pyogenes* strains ALAB49 and Xen20 cause systemic infections.

Indeed, since we now know that the *lux* operon incorporated into the Xen20 genome may be under the control of the *iolE* promoter, it is conceivable that the observed *in vivo* signal may be caused by a shift, rather than a decrease, in metabolic activity. Increasing utilization of *D*-glucose as the preferred carbon source may have led to a decrease of imaging signal below the detection threshold [107, 108]. It is also conceivable that this glucose-dependent promoter decreases the expression of *lux* operon-encoded proteins which convert endogenous fatty acids to usable substrates for the

luciferase enzyme, decreasing the overall pool of substrates, as previously suggested. Due to the observed effect of *D*-glucose on bioluminescence expression by Xen20 (Figure 2.5), further study is required to determine the specific effects of carbon source availability on regulation of bioluminescence expression and maintenance in *S. pyogenes* strains under carbon-limiting conditions. The effects of *D*-glucose on BLI signal by Xen20 *in vitro* strongly reflects the *in vivo* condition if one considers murine blood glucose ranges from ~70 to 90 mg/dl, which is ~3.9 to ~5 mM [109] and can be as high as 100 mg/dl (~ 5 mM) [110]. Consequently, the pathogen would be in an environment that is similar to the lower level of light detected in the *in vitro* time-course (Figure 2.5B and 2.5C). Therefore, in Chapter 3 we will investigate these observations in vitro using known metabolic models and various metabolic sources.

3. Dependence of Bioluminescence Expression on Central Metabolites in *S. pyogenes*

Introduction

Streptococcus pyogenes infections can account for severely invasive diseases associated with high mortality rates [3]. Two of the most severe forms of these invasive Group A Streptococcal (GAS) diseases are necrotizing fasciitis or the so-called "flesh-eating disease" and streptococcal toxic shock syndrome (STSS). For the 2013-provisional period, the national estimates reported by the Centers for Disease Control were expected to be at 11,500 cases projected based on 33 million individuals from the 10 reporting states, which were compiled for the current Active Bacterial Core Surveillance Report [111]. Of these GAS cases, 6-7% were necrotizing fasciitis or STSS, resulting in 20% and 50% mortality, respectively [111]. Early detection of invasive GAS infections and their discrimination from other casual pathogens is paramount to improving treatment efficacy and the better understand how *S. pyogenes* is capable to invade host immune defenses in the establishment of systemic disease [19, 75]. Given this, a distinct effort has been made to develop novel methods of tracking GAS *in vivo* through the use of molecular imaging [45]. More specifically, some imaging modalities report on total bacterial load [47], anti-microbial therapy [48], host immune response [112], and the relative location of both primary and distal infection sites [112]. As such, pre-clinical imaging of light-producing GAS has aided in tracking this dangerous pathogen *in vivo*. Biolumenscence in these engineered pathogens is most commonly driven by over-expression of either the *Photorhabdus luminescens lux* operon (*luxABCDE*) or the firefly luciferase (*ffluc*) plasmid followed by intraperitoneal luciferin injections minutes prior to imaging. We have previously described the use of *luxABCDE*-driven *S. pyogenes* Xen20 in the study of bacterial dissemination in murine models, which are deficient in the canonical streptokinase-plasminogen interaction (Chapter 2). Despite the lack of bioluminescent signal from distal organs in this model, histological and microbiological approaches determined the presence

of the pathogen, indicating a breakdown in the light production *in vivo* by Xen20. Using *in vitro* tests, we observed that exogenous *D*-glucose addition could alter light production by this organism. Here we aimed to probe this phenomenon further by altering the environment provided to the pathogen and measure the resultant light elaborated by the *luxABCDE* cassette-integrated pathogen in response to these triggers.

Bacterial bioluminescence driven by the *luxABCDE* operon initiates through the conversion of activated fatty acyl compounds into fatty aldehydes via the *luxCDE* system [113]. These fatty aldehydes are then passed to the *luxAB* complex, which utilizes molecular oxygen and $FMNH_2$ to produce the respective fatty acid, H_2O, FMN, and light. Therefore, one particular nuance of bioluminescent imaging using these expressed bacterial luciferases is the availability of necessary components of the luciferase system. First, the production of substrate for the luciferase enzyme is based on the catabolite levels in stationary-phase growth, where most bacteria exist within the infectious and environmental niche. Previous research by others [114] has shown recombinant *luxAB* genes from *Vibrio harveyi*, *V. fischeri*, and *Xenorhabdus luminescens* expressed in *E. coli* prefer aldehydes derived from heptanoic, octanoic, nonanoic, decanoic, and undecanoic acid, with the greatest luminescence observed from nonanoic acid. Similarly, the availability of reducing equivalent $FMNH_2$ and ATP are also related to the metabolism of the organism in the tissue of interest. Lastly, the availability of molecular oxygen may fluctuate greatly depending on the site of the infection, and may therefore have significant effects on the threshold of detection during *in vivo* models. Within, we outline the connection of catabolite repression of *S. pyogenes* Xen20 to bioluminescence production. In order to accurately reflect the bacterial load and the host-pathogen interface using bioluminescence expression and co-localization imaging studies, we must first understand the connection of light production (and, more specifically, the *de novo* synthesis of

luminescent substrates) to the means by which the bacteria responds to environmental or nutritional signals *in vivo*. In addition, adaptations of these techniques may be used as a sensor for the nutritional requirements and physiologic processes occurring in the pathogen.

Materials and Methods

Reagents, Chemicals, and Strains

Strains used for this study were purchased from PerkinElmer Inc. (Waltham, MA). Kanamycin was obtained from Research Products International (RPI; Mt. Prospect, IL). M9 minimal medium was produced as previously described [101]. Phosphate-buffered saline (PBS) concentrate (20x) was acquired from Teknova (Hollister, CA) and diluted to 1x with Milli-Q water before sterile filtering with a 0.22 µM syringe or autoclaving. In order to investigate carbon sources, we used the following additives: *D*-glucose (RPI), glycine (AMERSCO, Solon, OH), ATP (Sigma Aldrich, St. Louis, MO), AMP (Sigma Aldrich). Brain-Heart Infusion (BHI), Sheep Blood Agar (SBA) and Todd-Hewitt with yeast extract (THY) were obtained from BD Biosciences (San Jose, CA) or RPI. Unless otherwise specified, all reagents were acquired from Sigma Aldrich.

Bioluminescence Plate Assay

Strain Xen20 was streaked for isolation on one of either THY plates containing 0 or 50 mM *D*-glucose or SBA medium. Plates were grown at 37 °C for 20 hours and imaged using 60-second exposure, medium binning, and automatic aperture settings on an IVIS Lumina XRMS system (PerkinElmer Inc.).

Bioluminescence Kinetic Assays

Characterizing the bacterial virulence factors which exacerbate infections requires the ability correlate disease progression non-invasively as a surrogate for overall bacterial burden. As a first step, we chose here to use Xen20, a bioluminescent strain of *S. pyogenes* for establishment of the

foundation, as well as *Staphylococcus aureus* Xen29. Specifically, we used cells of Xen20 in stationary phase because this phase of growth was chosen as the majority of infections begin and persist with cells in stationary phase from the natural environment [115-118]. Strains Xen20 and Xen29 were grown in BHI broth for 18 hrs. at 37 °C in the presence of kanamycin (200 μg/mL). The cells utilized were in stationary phase, as confirmed by an optical density measurement $OD_{600nm} > 1$ which did not fluctuate; an OD_{600nm} of 1 unit of absorbance corresponds to 1.6×10^8 CFU per mL for Xen20 and 2.9×10^8 CFU per mL for Xen29, based on the strain technical information supplied by the vendor. The number of cells in each sample was adjusted to be equivalent prior to each assay. As previous data showed the *lux* cassette inserted under the *iolE* promoter [46], we first tested the contribution of *myo*-inositol as a transcriptional activator and glucose as a transcriptional repressor, based on analysis of this locus in other bacterial species [104, 119, 120]. Samples supplemented with *myo*-inositol or *D*-glucose were brought to a final volume of 1 mL with PBS. Samples were placed in the upright position within the IVIS Lumina XRMS on a 37 °C stage and imaged with a kinetic protocol. The imager was set to automatically adjust exposure times at 20-minute time points, for a total of 9 time points spanning 160 minutes. The flux (photons/second) of each sample was obtained from region of interest (ROI) for each sample using LivingImage software version 4.4 (PerkinElmer Inc.).

To assess the effect of carbon source availability on bioluminescence production, cells were combined with either filter-sterilized PBS (control) or PBS supplemented with 3 mM or 6 mM of the following: *D*-glucose or glycine. Imaging was repeated as before. In order to determine the contribution of growth rate to Xen20 expression, this process was then repeated for M9 minimal medium supplemented with 1% casein hydrolysate and 0.3% yeast extract.

For assessment of the relative effect that ATP has on the bioluminescence production, 1 mL of Xen20 cells was centrifuged at 14,800 rpm for 10 mins, the supernatant removed by pipet, and the cell pellet re-suspended in M9 minimal medium supplemented with ATP (10 or 100 µM) and AMP (10 or 100 µM), in the presence or absence of 3 mM *D*-glucose. Imaging was repeated as before, for a total of 140 minutes.

To mimic conditions similar to those found *in vivo*, cells were suspended in either: BHI broth; BHI broth supplemented with 5% sheep blood (SB); BHI broth supplemented with 3 mM *D*-glucose; BHI broth supplemented with 5% SB and 3 mM *D*-glucose; PBS supplemented with 5% SB and 3 mM *D*-glucose. The imaging and analysis protocol described above was repeated (8 total time points).

Luciferase Protein Assay

Xen20 cells were normalized for cell density (OD_{600nm}) and combined with PBS alone or PBS containing 3 or 6 mM of one of either *D*-glucose or glycine. Cells were incubated for 160 minutes before freezing in liquid nitrogen. The cells were thawed at 37 °C, and this process was repeated for a total of three freeze-thaw cycles. Cell lysates were treated with 100 µM of nonanal (Sigma Aldrich) using an auto-injector, previously shown to be a preferred substrate for the LuxAB luciferase [114]. Luminescence levels were measured in quadruplet on a Thermo Scientific VarioSkan plate reader.

Oxamate Inhibition Assay

In order to determine the contribution of L-lactate dehydrogenase to the luminescence of Xen20, Xen20 cells in stationary phase were added to PBS containing a final concentration of 6 mM *D*-glucose in a 96-well plate. Amounts of oxamate were added to a final concentration of 0-256 mM.

OD_{600nm} and luminescence were measured using a Thermo Fischer VarioSkan plate at 20 mins intervals over for 160 mins at a constant 37 °C.

Statistical Analysis

Statistical means of each group were analyzed using two-way, repeated measures analysis of variance (ANOVA) on R 3.1.2 (http://www.r-statistics.com/on/r/). Time, Substrate, and the interaction of Time*Substrate were analyzed for each medium and each bacterium. Results were reported for statistical values $p < 0.05$, or no significant difference for $p > 0.05$.

Results

D-Glucose Inhibits Bioluminescence from Xen20

Our previous results indicated the luminescence of Xen20 did not accurately reflect bacterial load *in vivo*, and suggested this may be due to a *D*-glucose-mediated inhibitory effects [121]. To confirm this finding on static medium, we plated Xen20 on THY plates supplemented with 0 or 50 mM *D*-glucose.

For this experiment, cells were grown overnight and imaged for luminescence (Figure 3.1). Colonies showed decreased luminescence for *D*-glucose supplementation. Interestingly, the presence of 5% sheep blood (SBA) also significantly decreased the bioluminescence produced by Xen20.

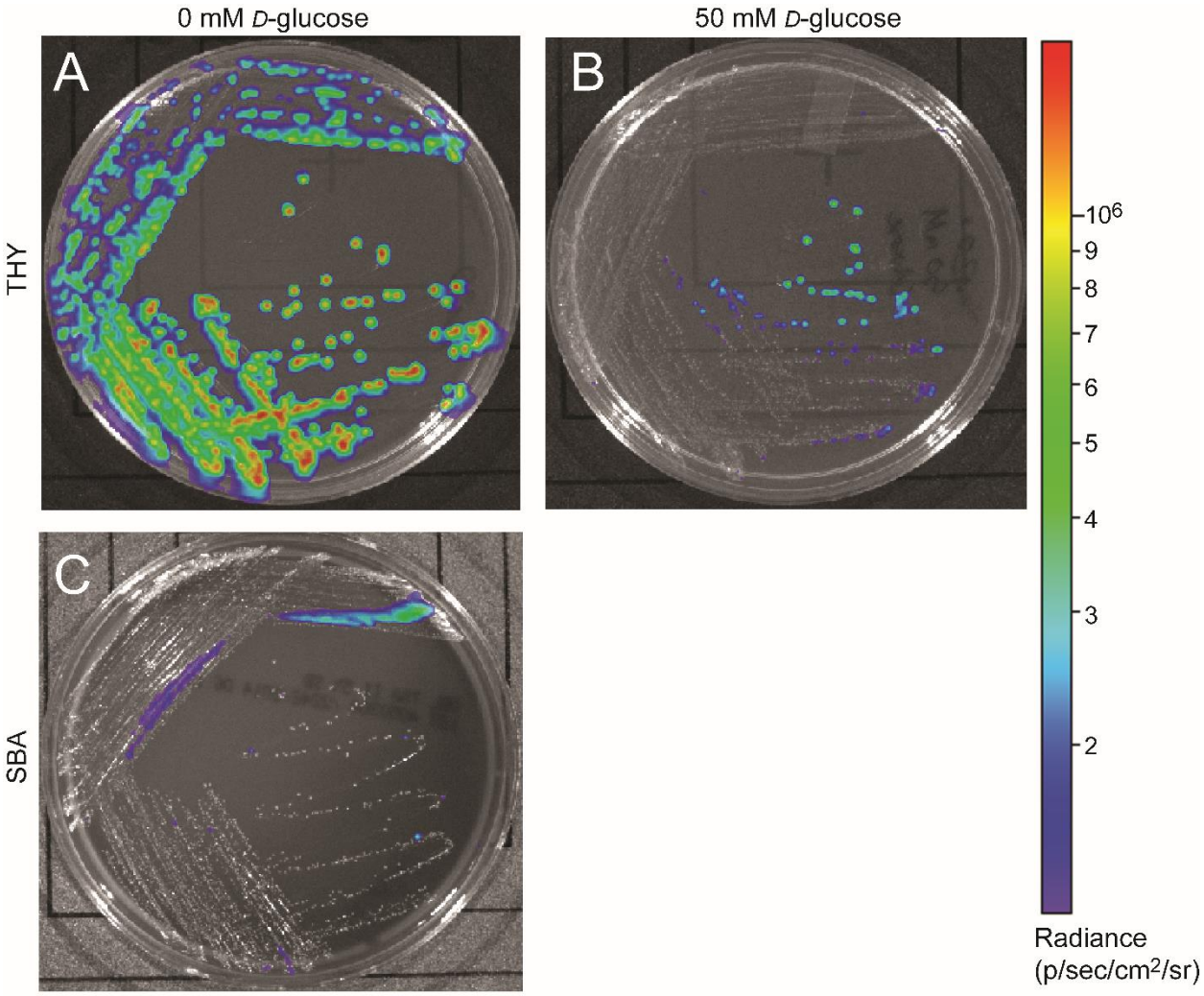

Figure 3.1: Glucose Effect on Bioluminescent Production of *S. pyogenes* Xen 20.

Xen20 bioluminescence on Todd Hewitt with yeast extract (*THY*) after 20 hrs. incubation at 37 °C in the without the addition of excess *D*-glucose (Panel A) or with the addition of 50 mM *D*-glucose (Panel B). Panel C shows similar treated Xen20 on sheep blood agar (*SBA*) without *D*-glucose supplementation.

The lux Operon Is Regulated by Inositol and Glucose Levels

The *luxABCDE* cassette is placed on a transposable element, and therefore expression is dependent on its random insertion into the genome of the bacterium [45]. In the case of *S. pyogenes* Xen20, the luxABCDE insertion was located within the *iol* operon, which has been shown in such species as *Bacillus subtilis* to encode enzymes for inositol catabolism, a substrate which is abundant in soil [119]. More specifically, the gene has inserted under the *iolE* promoter [46], which encodes 2-keto-*myo*-inositol dehydratase [119]. The entire *iol* operon is controlled by the regulator IolR [120]. Under conditions of high inositol, this repressor is released from the operator site, and the genes are expressed. In contrast, research has previously shown repression of green fluorescent protein (GFP) expression under the *iolE* promoter in *Salmonella enterica* serovar Typhimurium strain 14028 in glucose-rich conditions [104]. Therefore, we tested the contribution of these substrates to bioluminescence expression in Xen20. Exogenous glucose and/or inositol were provided as a sole carbon source in stationary-phase bacterial subcultures. The results (Figure 3.2) indicate increased luminescence from bacterial cells grown in the presence of *myo*-inositol. In contrast, co-administration of inositol with low concentrations of glucose (3 mM) reversed this effect.

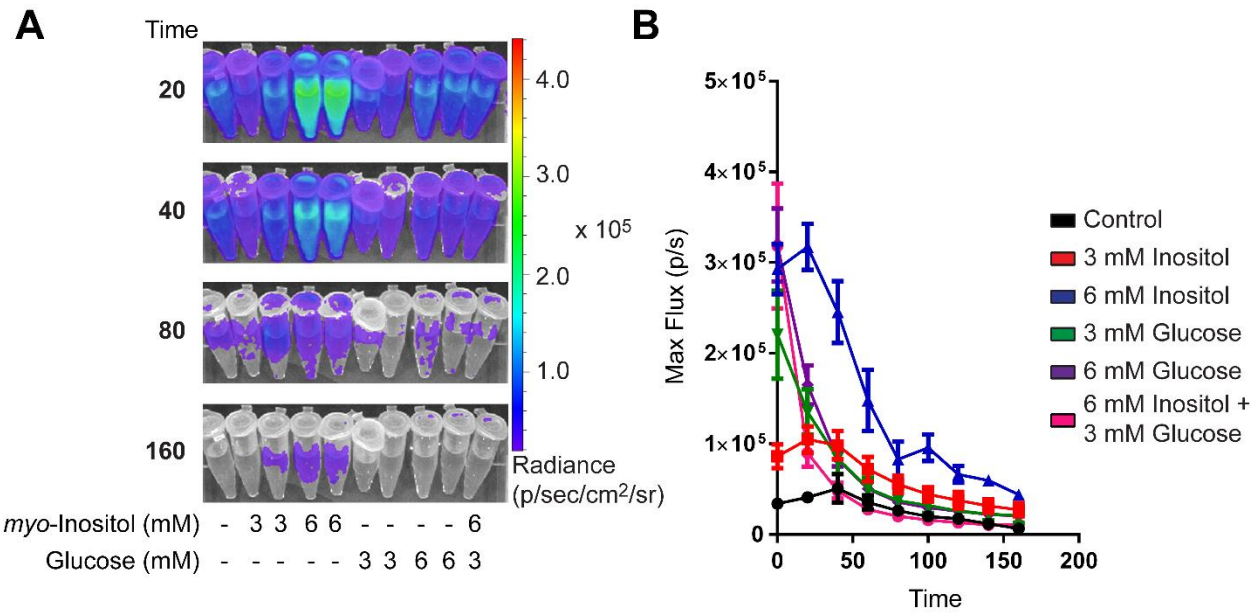

Figure 3.2: *Streptococcus pyogenes* Xen20 luminescence is inositol-dependent

A. Representative time course of Xen20 (*S. pyogenes*). For all samples, cells were grown to stationary phase then diluted in fresh phosphate buffered saline supplemented with either *myo*-inositol or glucose as indicated. *B* Quantification of luminescent signal from samples shown in Panel A for Xen20 pathogens represented as either the average rate of photons generated (shown as *Mverage flux*). Experiment performed as indicated in the Materials and Methods section.

Substrate Sources for Streptococcus pyogenes Bioluminescence

Since inositol cannot function as a carbon source in *Streptococcus pyogenes,* based on analysis of the KEGG genomic model (data not shown), we decided to investigate the synergistic or antagonistic effects of central metabolites to luminescence. As previously stated, bacterial bioluminescence is dependent on the production of substrates by the bacterial cell [113]. Therefore, we investigated the link of bioluminescence to central metabolism in *S. pyogenes* (see pathway scheme in Figure 3.3). Activated acyl donors, produced *de novo* by the bacterial cell, serve as the substrate for the *lux* operon to create fatty aldehydes, which are then converted to fatty acids in the presence of $FMNH_2$ and oxygen, releasing light. Therefore, the bioluminescence is dependent entirely on the availability of these activated acyl donors. These donors can be produced in two ways: first, coenzyme A (CoA)-containing acyl compounds can be created via the breakdown of fatty acids by β-oxidation (yellow box, Figure 3.3); second, CoA-containing fatty acid building blocks can be created via the fatty acid biosynthesis pathway (red box, Figure 3.3). This second pathway is dependent upon the creation of acetyl-CoA by the glycolysis/lactic acid fermentation pathway (blue box, Figure 3.3).

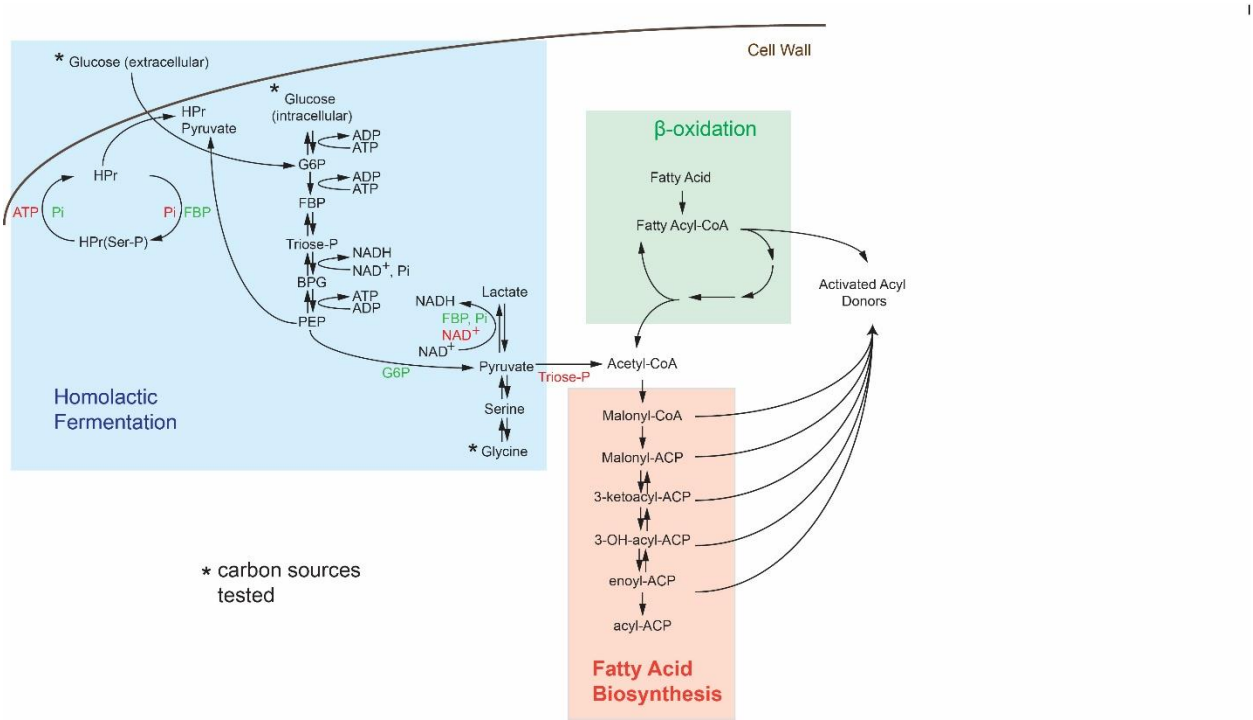

Figure 3.3: *De novo* synthesis of acyl donors via central carbon metabolism in *Streptococcus*. Activated acyl donors, the substrates for the *luxABCDE* machinery, are made by breakdown of fatty acids through *β*-oxidation (*green region*) or by the synthesis of activated carrier protein (*ACP*)-containing metabolites via fatty acid biosynthesis (*salmon region*). Since *S. pyogenes* lacks the *β*-oxidation pathway, all substrates for light production in Xen20 are created by acetyl-coenzyme A (*CoA*)-dependent fatty acid biosynthesis produced by homolactic fermentation (*blue box*). Asterisk indicate sole carbon sources tested. Allosteric activators, such as inorganic phosphate (*Pi*), fructose bisphosphate (*FBP*), and glucose-6-phosphate (*G6P*) are shown in green text and allosteric inhibitors, namely adenosine triphosphate (*ATP*), nicotinamide adenine dinucleotide (NAD^+), and triose phosphate (*Triose-P*) are shown in *red text*. Abbreviation used in here are as following: histidine-containing carrier protein (*HPr*), phosphorylated HPr (*HPr(Ser-P*), phosphoenolpyruvate (*PEP*), and 1,3-bisphosphoglyceric acid (*BPG*).

Although Kypto Encyclopedia of Genes and Genomes (KEGG) Pathways are not available for *S. pyogenes* Xen20, a pathway does exist for the closely related M49 serotype ancestor NZ131 (http://www.genome.jp/dbget-bin/www_bget?genome:T00780). Analysis of this pathway revealed important allosteric and feedback mechanisms for bioluminescence production. First, M49 strains lack the *β*-oxidation pathway beyond the creation of hexadecanoyl-CoA from hexadeconate (Figure 3.4). In contrast, *S. aureus* NCTC8325, the parental strain of a frequently manipulated strain RN4220, showed increased capacity for degradation of fatty acids (Figure 3.4). Therefore, the majority of activated acyl donors must be created in these M49 strains by the fatty acid biosynthesis pathway (dependent on the creation of acetyl-CoA by glycolysis). Analysis of these pathways revealed M49 to have all necessary enzymes for these pathways (data not shown). Therefore, we decided to further investigate the contribution of glycolytic compounds on bioluminescent signal.

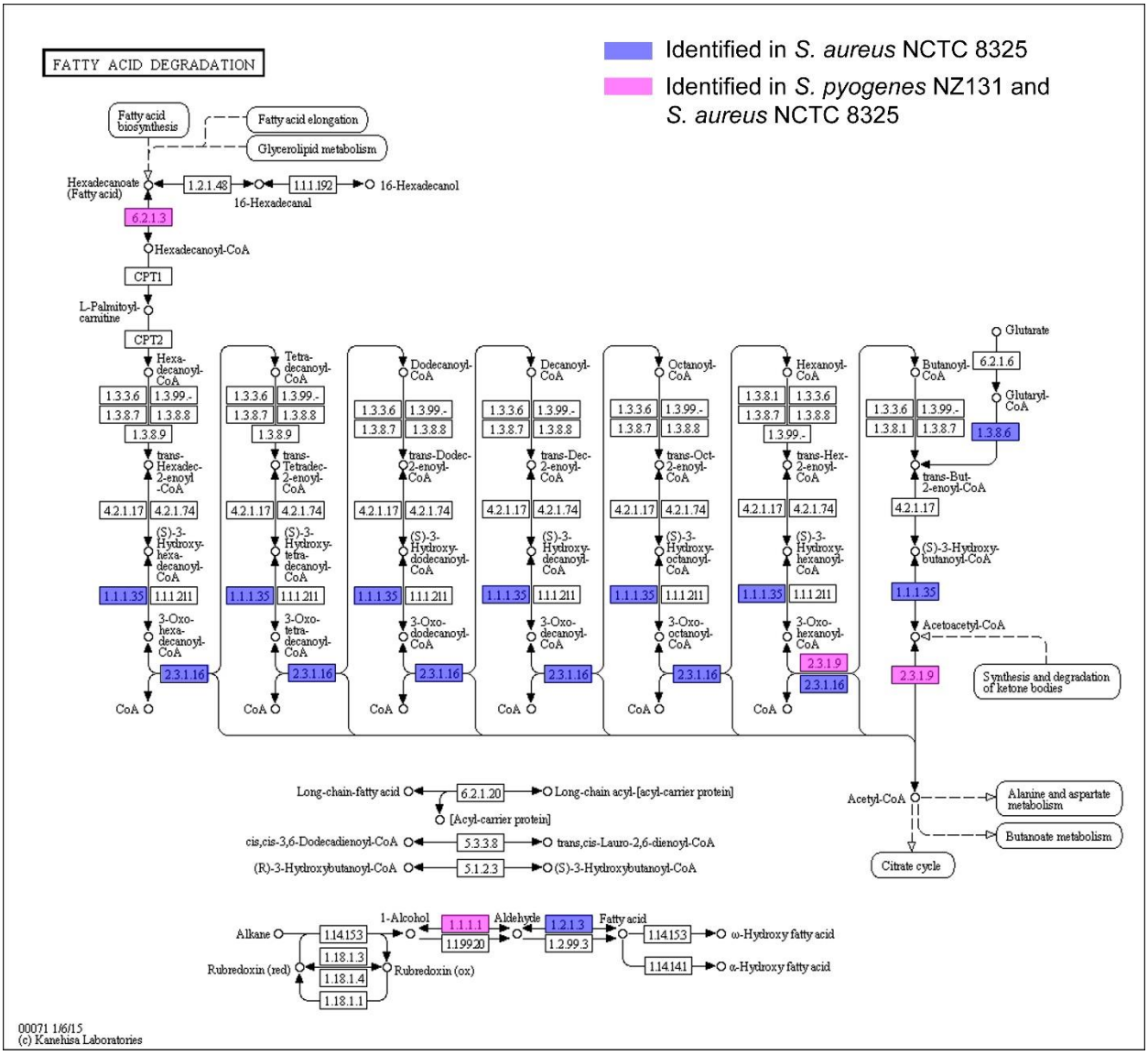

Figure 3.4: KEGG Genome Analysis of *S. pyogenes* and *S. aureus* β-oxidation Pathways.

Shown is the KEGG genome reference pathway for β-oxidation (map0071). Enzymes identified in the genomes of *S. aureus* NCTC 8325 (*blue boxes*) and in both *S. aureus* NCTC8325 and *S. pyogenes* NZ131 (*purple boxes*) are indicated by enzyme commission numbers. No enzymes were identified in *S. pyogenes* NZ131 alone.

Dependence of Luminescence on Glucose Homeostasis

S. pyogenes (Xen20) processes sugars by homolactic fermentation, a process which utilizes *D*-glucose as the preferred carbon source [122]. In contrast, based on the results of KEGG genome analysis, glycine is converted to pyruvate via conversion to serine, and production of compounds upstream of pyruvate is dependent on the gluconeogenesis pathway. *D*-Glucose is present at physiological concentrations ranging between 3.9-5 mM in a healthy adult mice [109]. Therefore, exogenous *D*-glucose and glycine were provided at a concentration of either 3 or 6 mM. In order to contrast the results from this lactic acid bacterium, *Staphylococcus aureus* Xen29 was included as a comparative control, as it is a bioluminescent, respiring pathogen. Bioluminescent imaging at sequential time points revealed addition of any substrate to Xen20 decreased the bioluminescence expression in stationary phase of the cells (Figure 3.5A). Counter to this, addition of *D*-glucose or glycine increased luminescence expression in *S. aureus* Xen29 (Figure 3.5B). Subsequent serial dilutions and plating of cells in these tubes indicated equivalent CFU and bioluminescence recovery after 18 hours of plate growth, indicating this change in bioluminescence is not due to fluctuations in cell number or *lux* instability (data not shown).

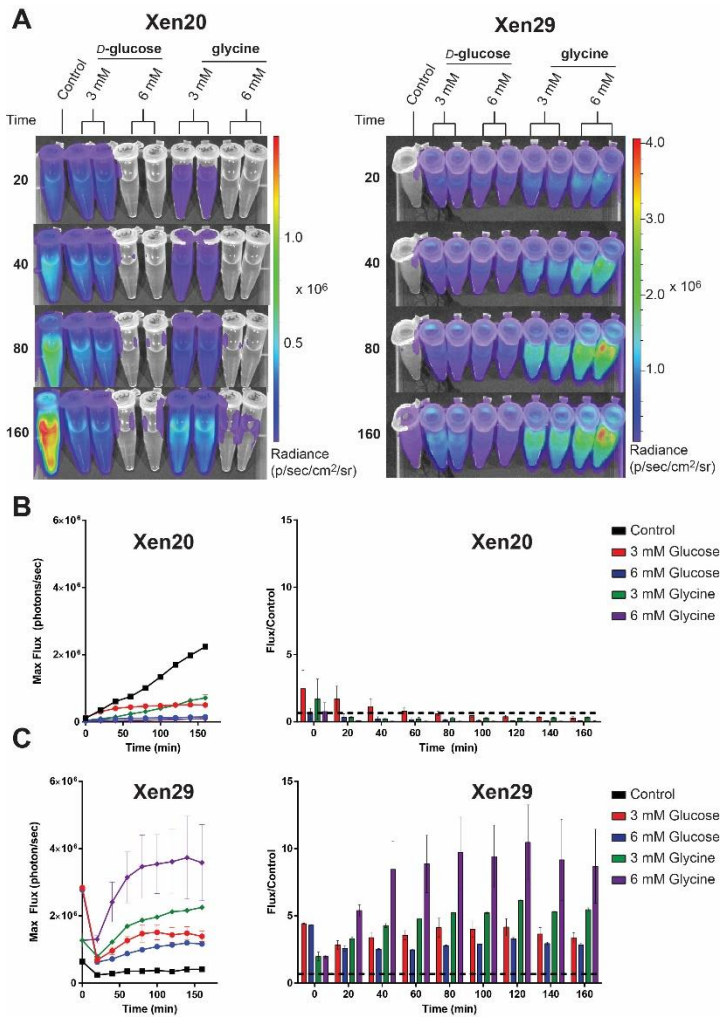

Figure 3.5: Evidence of bioluminescent colonies retrieved from *in vivo* infection.

A. Representative time course of Xen20 (*S. pyogenes*) or Xen29 (*S. aureus*). For all samples, cells were grown to stationary phase then diluted in fresh phosphate buffered saline (*Control*) supplemented with either *D*-glucose or glycine as indicated. *B-C.* Quantification of luminescent signal from samples shown in Panel A for both Xen20 and Xen29 pathogens represented as either the maximum rate of photons generated (shown on the *left* as *Max flux*) or as the relative maximum signal normalized to the control (shown on the *right* as *Flux/Control*) displayed as a function of time. Black dashed line indicates a ratio at threshold of 1.0 and experiment performed as indicated in the Materials and Methods section.

This produced luminescence signal may be multi-faceted and interpretation of it complicated by the overall health of the cell or permeability of the necessary substrates acquired from the local environment. In order to test the latter, we perturbed the carbon source equilibrium of both light producing pathogens (i.e., Xen20 and Xen29) by characterizing the glucose-dependency of bioluminescence expression starting from relatively minimal *D*-glucose levels (e.g., M9 minimal) to incrementally higher *D*-glucose levels (Figure 3.6A). To maintain the overall viability of these cells as far as the nascent metabolic needs of *S. pyogenes* and *S. aureus*, their respective medium was also supplemented with both casein hydrolysate and yeast extract to provide amino acids and cofactors for protein and DNA synthesis. In this moderately enhanced M9 medium, we saw that supplementation with 3 and 6 mM *D*-glucose increased overall the bioluminescent signal for Xen20 until it intersects the control level at 140 minutes leading to an ultimate ~5-fold reduced light level at 160 minutes compared to control (red /blue lines with glucose and black lines for controls shown in Figure 3.6B). This production of light changes was compared to samples alternatively supplemented with glycine (3 mM), which showed a comparative decrease in signal compared to *D*-glucose levels at 100 minutes but recovered to above *D*-glucose levels at 160 minutes. This was overcome more rapidly with the higher glycine dose (6 mM) that exhibited an increased signal, which plateaued at 100 minutes and dipped below the 3 mM glycine light production level at 160 minutes. In contrast, all supplementations increased bioluminescent expression in Xen29 cells (Figure 3.6C).

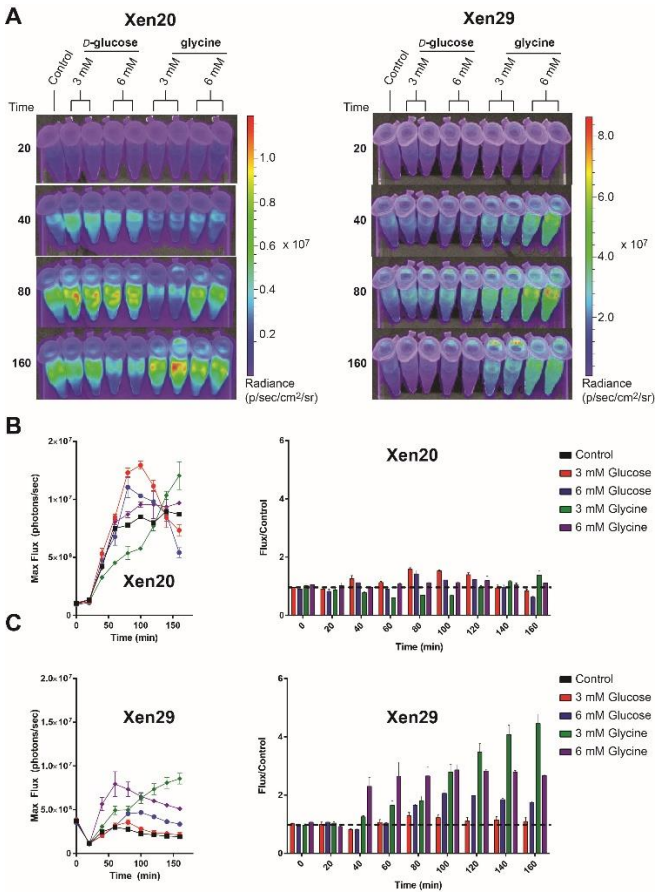

Figure 3.6: Increased luminescence production by Gram-positive pathogen incubated in M9 medium supplemented with casein hydrolysate and yeast extract.

A. Representative time course of Xen20 (*S. pyogenes*) or Xen29 (*S. aureus*). For all samples, cells were grown to stationary phase then diluted in fresh M9 medium containing casein hydrolysate and yeast extract (*Control*) supplemented with either *D*-glucose or glycine as indicated. *B-C.* Quantification of luminescent signal from samples shown in Panel A for both Xen20 and Xen29 pathogens represented as either the maximum rate of photons generated (shown on the *left* as *Max flux*) or as the relative maximum signal normalized to the control (shown on the *right* as *Flux/Control*) displayed as a function of time. Black dashed line indicates a ratio at threshold of 1.0 and experiment performed as indicated in the Materials and Methods section.

The maximum flux for each supplementation and medium was prepared and analyzed using two-way, repeated measures ANOVA to determine the effects of three variables on bioluminescence: 1) time, which determines the effect of time on the bioluminescence in that particular supplemented medium; 2) substrate, which determines the effect of differing substrates at each time point; and 3) interaction, which determines the significance of the interaction of these two variables on bioluminescent expression (Table 3.1). The results indicate that *D*-glucose and glycine statistically affect the bioluminescence at all levels of supplementation, both individually and in interaction with time. However, this interaction was shown to not be statistically relevant upon addition of the casein hydrolysate, yeast extract, and M9 salts. In contrast, Xen29 showed only significant interactions for glycine supplementation (at both 3 and 6 mM), in both PBS and M9 medium. *D*-Glucose was found to only be significant in PBS without interaction of time in Xen29.

Table 3.1: Two-Way Anova Results for Substrate-Dependent Luminescence

Medium	SUPPLEMENT	MEASURE	p-value
PBS	Glucose	Time	< 0.001
		Substrate	< 0.001
		Time*Substrate	< 0.001
	Glycine	Time	< 0.001
		Substrate	< 0.001
		Time*Substrate	< 0.001
M9	Glucose	Time	< 0.001
		Substrate	0.634
		Time*Substrate	0.964
	Glycine	Time	< 0.001
		Substrate	0.4683
		Time*Substrate	0.3393

Dependence of Energetics on Luminescence Expression

In addition to glucose homeostasis, bioluminescence is dependent on the availability of cellular ATP. As outlined earlier, this energy is used during the redox reaction, which converts $FADH_2$, O_2, and fatty aldehyde substrates to fatty acids, H_2O, and FMN, producing light. ATP is expended during the initial conversion of fatty acids to their respective fatty aldehydes by the LuxCDE complex [123], as well as providing several allosteric inhibitions in the glycolytic pathways (Figure 3.3). Therefore, we assayed the effect of supplementing M9 with 10 μM or 100 μM of ATP or AMP. The addition of ATP and AMP did not stimulate the production of bioluminescence in cells (Figure 3.7). Interestingly, the addition of 3 mM *D*-glucose recovered the luminescence of these cells, although this level of luminescence was still below that previously observed for cells in 3 mM glucose in M9 medium (Figure 3.6B).

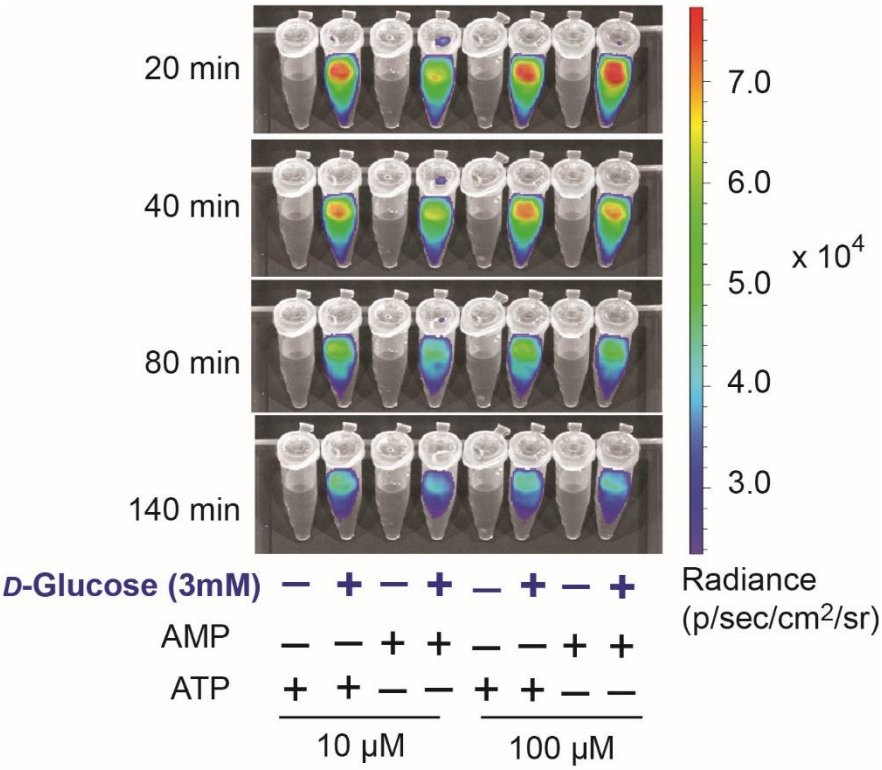

Figure 3.7: Partial recovery of luminescence signal upon addition of *D*-glucose and ATP or AMP, but not ATP or AMP alone

Xen20 (*S. pyogenes*) cells in stationary phase were washed and resuspended in M9 medium and combined with the indicated concentration of ATP or AMP with and without with 3 mM *D*-glucose supplementation.

Inhibition of L-lactate dehydrogenase

Pyruvate created by the homolactic fermentation is either converted to acetyl-CoA by pyruvate dehydrogenase or to lactate by L-lactate dehydrogenase. In order to test the effects of pyruvate conversion to lactate in the availability of the acetyl-CoA pool, Xen20 cells were combined with oxamate, an inhibitor of L-lactate dehydrogenase. Cells were combined with amounts of oxamate ranging from 0-64 mM, in the presence of 6 mM glucose in PBS. At levels of oxamate above 32 mM, growth of the cells was inhibited, indicated by a decrease in the OD_{600nm} over the observed time (Figure 3.8A). However, at concentrations below this threshold, no significant difference was observed in luminescence (Figure 3.8B). All observed levels were below that observed for cells in PBS medium alone (Figure 3.8B, black line).

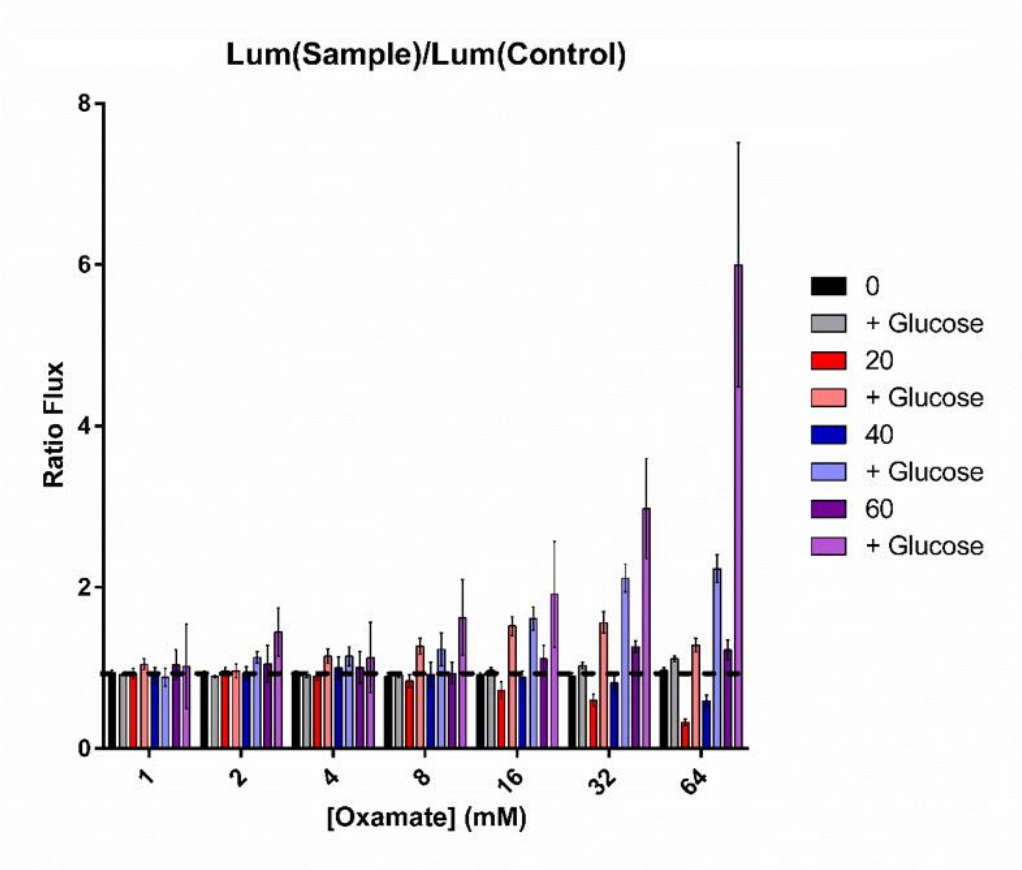

Figure 3.8: Oxamate inhibition does not recover luminescence signal.

Xen20 cells in stationary phase were combined with concentrations of PBS alone or PBS with 6 mM glucose containing 0-64 mM oxamate. Graphed are the summation of eight trials. Shown is the quantification of the ratio of average flux of cells treated with oxamate versus untreated (*Ratio Flux*) as a result of oxamate concentration. Colors represent different time points with or without glucose concentrations. Black dashed line indicates a ratio at threshold of 1.0 and experiment performed as indicated in the Materials and Methods section.

In vitro Model of Tissue Luminescence

Finally, in an attempt to mimic *in vivo* conditions, Xen20 cells were combined with distinct media combination containing either BHI broth, BHI broth supplemented with 5% SB, BHI broth supplemented with 3 mM *D*-glucose, BHI broth supplemented with 5% SB and 3 mM *D*-glucose, and PBS supplemented with 5% SB and 3 mM *D*-glucose (Figure 3.9). The samples containing SB exhibited a steady decrease in bioluminescence, irrespective of *D*-glucose concentration or presence of BH; this result may be attributed to the light absorptive properties of hemoglobin, which may also have a potential concomitant interference on such *in vivo* BLI studies. In a comparison of BHI broth to BHI broth supplemented with 3 mM *D*-glucose, supplementation with *D*-glucose produced a greater decrease in bioluminescence, despite similarities in initial bioluminescence.

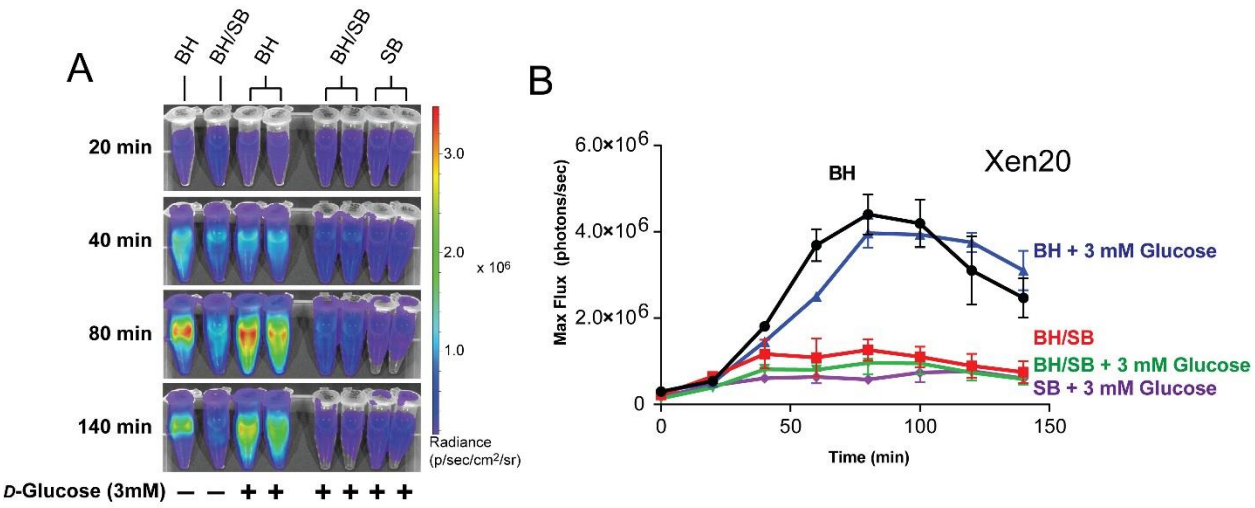

Figure 3.9: Effect of blood supplementation of biolumenescence signal propagation by *S. pyogenes* Xen20.

A. Representative time course of Xen20 (*S. pyogenes*). Cells in stationary phase were diluted in brain heart infusion broth (*BH*) and/or 5% sheep blood (*SB, BH/SB*) in the presence or absence of with *D*-glucose to a final [3mM]. *B.* Quantification of luminescent signal from samples shown in Panel A for the Xen20 pathogen represented as the maximum rate of photons generated (*Max flux*).

Discussion

Competition for carbon sources and energy may have a significant effect the survival of *Streptococcus pyogenes in vivo* during an infection and, as such, we explored the use of a bioluminescent strain of *S. pyogenes* that exhibited an accruement for sensing these environmental factors as manifested altered bioluminescence production. Our hypothesis was supported by the initial *in vitro* observation that this stably incorporated *luxABCDE* pathogen Xen20 displayed reduced light production in *D*-glucose enriched medium. Here we expanded the score of these findings to characterize the dose dependent responsivity of this glucose phenomenon and to test the ability of Xen20 to show changes when other triggers are provided (e.g., addition of ATP). Interesting, our results indicate that addition of exogenous ATP completely blocked light production in our integration time measured by the detector but that providing additional carbon sources as fuel for the microbe counteracted this effect.

Originally, genetic evidence pointed to *iolE* as the gene sequence interrupted by the *luxABCDE* pathway [46]. This gene encodes the 2-keto-myo-inositol dehydrogenase, involved in the inositol phosphate pathway in strains such as *Bacillus subtilis* [119]. Therefore, the results of inositol supplementation (Figure 3.2) indicate a genetic regulation of the *luxABCDE* bioluminescence. According to the KEGG metabolic pathways for NZ131 (a related member of the M49 serotype), Streptococcus pyogenes does not contain the enzymes necessary to convert this inositol further into usable carbon sources to enter glycolysis or citric acid cycle. Therefore, we decided to also test the effects of other carbon sources in order to link the creation of luminescent substrates to central metabolism of the bacterial cell.

Our analysis of the KEGG metabolic pathways for NZ131 revealed no enzymes capable of fatty acid degradation beyond the creation of hexadecanoyl-coenzyme A (see the scheme in Figure 3.3 and the KEGG pathway in Figure 3.4). Therefore, activated fatty acyl compounds must be created

via the fatty acid biosynthesis pathway, which utilizes acetyl-CoA as building blocks for the creation of fatty acid chains. As previously mentioned, these compounds are produced in *S. pyogenes* via homolactic fermentation. Taken together this would suggest *S. pyogenes* has a simplified route to production of the essential light producing building block than its counterpart *S. aureus*, which was use here, in many cases, to contrast observed effects, thereby eliminating the possibility of the direct effects on the exogenous materials on the luminescence post-production by Lux operon.

Again for the sake of clarity, glucose is utilized in one of two ways: first, it can be transported into the cell by phosphotransferases, producing glucose-6-phosphate, which is fed into the fermentation process. In contrast, intracellular glucose is converted by hexokinase into glucose-6-phosphate [122]. Addition of *D*-glucose and glycine were shown here to decrease the bioluminescence of Xen20. As seen in Figure 3.3, and reviewed elsewhere [122], excess glucose during homolactic fermentation causes the conversion of pyruvate to lactate by the enzyme L-lactate dehydrogenase due to increased levels of fructose-bisphosphate (FBP), intensified by the inhibition of pyruvate conversion to acetyl-CoA by increased triose phosphate levels. Therefore, the decreased bioluminescence observed in *D*-glucose supplementation may be due to the decreased acetyl-CoA pool, and subsequent lack of fatty acid biosynthesis precursors. In order to test this hypothesis, we subjected cells in glucose supplementation to levels of oxamate, an inhibitor of L-lactate dehydrogenase, in order to restore the acetyl-CoA pool and, subsequently, luminescent signal. However, this inhibition proved to decrease the growth of cells *in vitro*, indicating the necessity of this pathway in Xen20 physiology (Figure 3.8). Interestingly, glycine supplementation showed comparable bioluminescence levels at concentrations equal to those of *D*-glucose. Although

glycine is first converted to pyruvate, gluconeogenesis can convert this pyruvate back to triose phosphate and/or FBP, creating a similar effect to *D*-glucose supplementation.

Interestingly, addition of amino acid sources, such as casein hydrolysate, relaxed the inhibition of bioluminescent signal by *D*-glucose and glycine supplementation (Figure 3.6). According to the KEGG diagram for *S. pyogenes* NZ131, L-alanine can be converted to pyruvate and would contribute to luminescence. In contrast, amino acids which are converted to fumarate, are unable to contribute, due to the lack of citrate cycle. Leucine, valine, and isoleucine are converted to their subsequent oxopentanoates, but subsequent reactions are not possible due to a lack of 3-methyl-2-oxobutanoate dehydrogenase. Therefore, the contribution of amino acid supplementation will require further study.

Diminished luminescent was also seen with ATP and AMP supplementation, which is counterintuitive given the dependence of fatty aldehyde production on ATP hydrolysis. Due to the similar activity of ATP and AMP, we hypothesize this observed effect is due to the transport of adenosine. The effects of low AMP and ATP concentrations (< 1 mM) on adenosine uptake in bacteria such as *Escherichia coli* have long been known [124]. However, the adenosine transporters of *S. pyogenes*, and their contribution to in vivo metabolism, have not been well described. In *E. coli*, ChIP-chip analysis revealed the PurR regulator to control the adenine transporter *yieG* [125]. Indeed, the PurR regulator in *Lactococcus lactis* controls expression of genes relevant to purine uptake, including *bmpA*, *pbuO*, and *pbuX*, and is dependent on the presence of absence of exogenous purine sources [126]. However, this system is little understood in *S. pyogenes*, and will require future study to link these regulatory elements and transporters to extracellular nucleoside concentrations.

If indeed the uptake of adenosine and subsequent conversions to ADP and ATP are inhibited, this would have a range of effects on the central metabolism of Xen20 and its link to light production. The change in levels of cellular ATP may have a direct effect on the phosphotransferase-dependent import of glucose through restricting the conversion of HPr(Ser-P) to HPr. ATP is also involved in two steps of the glycolysis pathway: first, in the conversion of intracellular glucose to glucose-6-phosphate (G6P); and second, G6P to fructose 1,6-bisphosphate. Previous data has shown that the uptake of extracellular inorganic phosphate (P_i) occurs through either an ATP-dependent active transporter or a concentration-dependent, facilitated diffusion transporter [13]. The passive transport of extracellular phosphate is inhibited by levels of ATP, focusing instead on the active transport system, which increases levels of ADP and P_i through hydrolysis of existing intracellular ATP. Increased levels of ADP and P_i modulates homolactic fermentation, including: 1) increasing production of G6P via the phosphotransferase system by increasing HPr(Ser-P) conversion to HPr, as well as inhibiting conversion of intracellular glucose to G6P; 2) providing additional cofactors (ADP) for the conversion of BPG to PEP and PEP to pyruvate, increasing the pyruvate pool; and 3) inhibiting the conversion of PEP to pyruvate and promoting the subsequent conversion of pyruvate to lactate. These first two effects would seemingly increase luminescent substrate production through increased glycolytic activity, and the third effect would decrease the substrates available for its production. Based on the decreased levels of bioluminescence detected in the presence of ATP (Figure 3.7), it seems that this inhibition of PEP conversion and promotion of lactate formation has a greater effect on bioluminescent signal than the promotion of upstream glycolytic reactions. Addition of *D*-glucose was shown to overcome this inhibition, although levels observed were still below that of Xen20 cells grown in the absence of exogenous ATP or

AMP. This could be due to a reactant-based increase in steady-state reactions, overcoming the decreased pyruvate (and subsequent acetyl-CoA) pool.

To our knowledge, these results are the first to highlight the effects of the central metabolism of the bacterial luciferase system with its subsequent light production. We recommend these engineering of these light-producing microbes must first characterize the effects of these metabolites on light production, as some like Xen20 are keenly more sensitive to external cues than Xen29 for instance. In cases where organisms are revealed to have allosteric inhibitions dependent on metabolite levels, luciferases, which are independent on the de novo creation of luciferase substrates, such as the *ffluc*, system previously described in *S. pyogenes* [61], may be considered, in which these luciferase substrates would be provided exogenously rather than made exclusively by the cell.

Application Note: Characterizing Antibiotic Mechanisms Using Bioluminescence Imaging

Antibiotic resistant strains of bacteria are increasing constantly in the clinical setting. Indeed, the CDC 2013 Threat report lists three bacteria with an Urgent Threat Level (*Clostridium difficile*, carbapenem-resistant *Enterobacteriaceae*, and Drug-resistant *Neisseria gonorrhea*), and 11 additional species with a threat level of Serious, including Methicillin-Resistant *Staphylococcus aureus* (MRSA) [2]. Therefore, there is a need for compounds with activity against these resistant strains. The discovery of these compounds can be time-consuming and costly, and it is no surprise that the report indicates the number of new antibiotics developed and approved has decreased significantly in the last three decades.

Previously, we have shown that physiological conditions of the bacterial cell are dependent on the expression of bioluminescence. As the reaction requires a combination of oxygen, reducing equivalents (NADH and $FMNH_2$), and luminescent substrates created from fatty acids, the

reduction of any of these factors can cause a decrease in bioluminescent signal. We have shown the effects of catabolite repression on the bioluminescence of *S. pyogenes* Xen20 (Chapter 3). Using this knowledge, we sought to design streamlined protocols for the characterization of antibiotics based on: 1) effective concentrations and 2) mechanisms of action.

In genetically modified bacterial luciferases, fatty acids are converted to fatty aldehydes by the LuxCDE complex. These activated acyl donors then serve as substrates for the light-producing steps of the luciferase reaction. By varying the medium conditions, we were able to differentiate antibiotic mechanisms in a Gram positive (*S. aureus* Xen29) and Gram negative (*E. coli* Xen14) bacterium. Cells were exposed to known bacteriostatic (levofloxacin and linezolid) and known bactericidal (meropenem and oxacillin) agents. Our results (Figure 5.1) indicate a "halo" effect which occurs at the threshold of bacterial growth in the presence of bacteriostatic agents, which does not occur in the presence of bactericidal agents. For *E. coli* Xen14, this differentiation was found to be uniquely significant for bacteriostatic agents in two mediums: eosin methylene blue (EMB), and MacConkey agar (MAC). These mediums are known to be rich in fatty acids, due to high peptone content. Similarly, *S. aureus* Xen29 was observed to be uniquely significant for bacteriostatic agents in brain heart infusion medium, a rich medium high in fatty acid content.

Interestingly, this shift in bioluminescent signal in sub-clinical levels of bacteriostatic agents was found to be due, at least in part, to levels of reducing equivalents. *S. aureus* Xen 29 cells treated with bactericidal agents showed a ten-fold increase in the ratio of reduced/oxidized equivalents, as opposed to those treated with bacteriostatic agents. As these oxidized

These observations illustrate one potential application of physiological bioluminescence dependence. Future study on this topic will be required to test unknown compounds for their antibiotic properties, followed by characterization through traditional means to validate the

approach. In addition, the robustness of this model needs to be validated using a wider range of bacterial pathogens, such as GAS, *Pseudomonas aeruginosa*, *Klebsiella pneumonia*, and others.

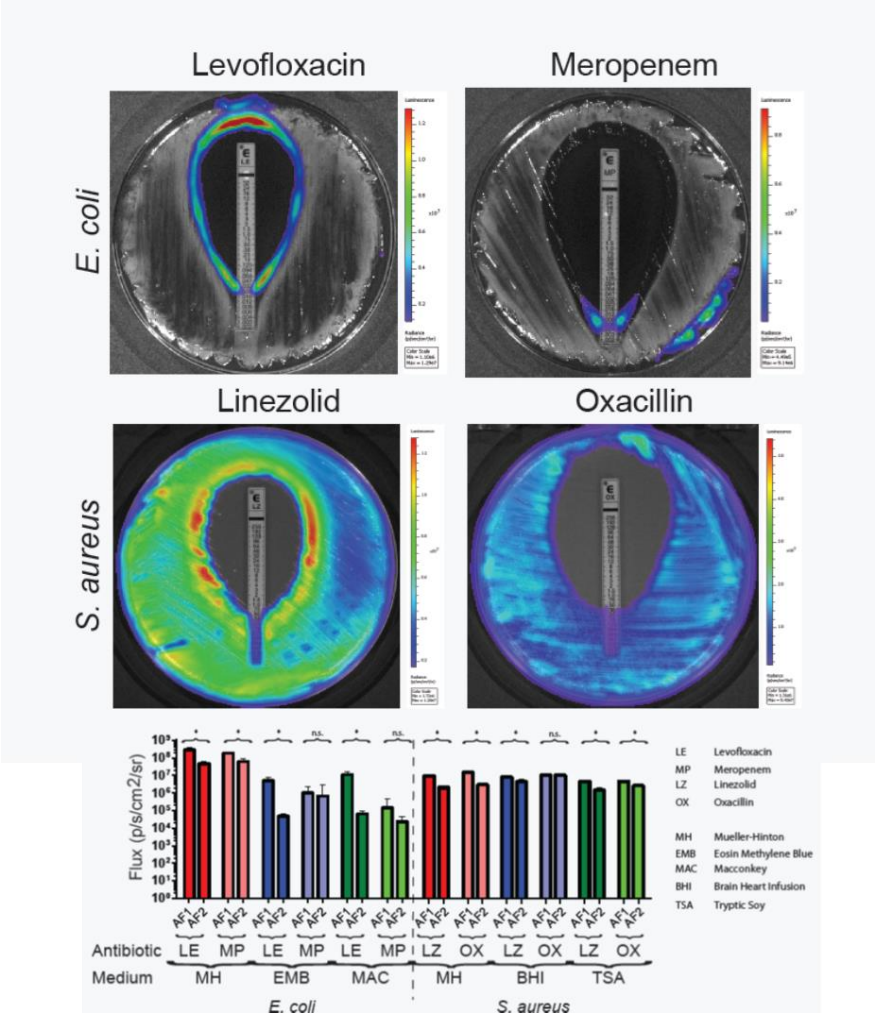

Figure 3.10: Medium Dependence of Lumincescence on Antimicrobial Differentation.

E. coli Xen14 and *S. aureus* Xen29 were grown for twenty hours in various medium compositions to determine the signal-to-noise (AF1 vs. AF2) ratio of luminescence in subclinical levels of bacteriostatic (levofloxacin and linezolid) or bactericidal (meropenem and oxacillin). Abbreviations: LE, levofloxacin; MP, meropenem; LZ, linezolid; OX, oxacillin; MH, Mueller-Hinton; EMB, eosin methylene blue; MAC, MacConkey; BHI, brain heart infusion; TSA, tryptic soy agar. * indicates $p < 0.05$, otherwise no significant difference (n.s.d.)

4. Genomic and Pathogenic Characterization of an Ancestral *Staphylococcus aureus* Strain Potent in Systemic Models of Infection

Introduction

The prevalence of methicillin-resistance in *Staphylococcus aureus* is becoming a serious, global threat as noted by the Centers for Disease Control and Prevention 2013 Threat Report [2]. *S. aureus* strains are heterogeneous, and therefore can cause a wide range of disease, from localized infections, such as impetigo and cellulitis, to systemic infections, such as bacterial endocarditis. In particular, *S. aureus* is the leading cause worldwide of bacterial endocarditis, a condition associated with mortality rates as high as 25-47%, even in the presence of antibiotic therapy [127]. This wide range of competence for infection, as well as the growing number of antibiotic resistance profiles, is due to the heterogeneity among *S. aureus* strains, driven primarily by the acquisition of mobile genetic elements (MGE's) [38], which comprise what is referred to as the "accessory genome." The most famous of these elements is the SCCmec operon, which encodes the *mecA* gene and confers methicillin resistance [38]. However, many other MGEs can confer adaptive advantages during infection, such as enterotoxins [128], leukocidins (most famously the Panton-Valentine leukocidins, *pvl*, encoded by the *lukF-PV* and *lukS-PV* genes) [128] [38], staphylokinase (*sak*) [129], and the toxic shock syndrome toxin TSST-1 [130]. Therefore, there is an urgent need to better understand the evolutionary adaptations commonly associated with acquisition of mobile genetic elements that contribute to (i) the prevalence of nosocomial infections, (ii) the rapid spread of community-acquired strains, (iii) development and expansion of novel resistance mechanisms, and (iv) natural host selection and propagation pathways.

Historically, primary disease isolates from patients with infections have long been phenotyped and categorized by their colony appearance on agar plates, expression of catalase, hemolytic potential, and ability to clot rabbit plasma [131]. As is central to these types of analyses, the clinical

distinction between *S. aureus* and *S. epidermidis* is dependent on coagulase activity assessed by observable clot formation [132], which demonstrates the presence of staphylocoagulase and von Willebrand factor binding protein (vWBp) [31, 32]. More recently, research is focusing on characterizing genomic markers for various infectivity and resistance profiles. For example, work is underway to design an allelic profile to indicate vancomycin-intermediate *S. aureus* (VISA), a phenotype that is the result of point mutations in several genes and displays an increased resistance to vancomycin treatment [54, 133]. By statistically analyzing the connection between these observed phenotypes and the presence of small nucleotide polymorphisms (SNPs) and/or MGE-encoded genes, this may identify particularly virulent strains and inform clinical care. However, these genome-wide association studies (GWAS) are limited by the need for a reference strain for SNP calling, and resort to references which may not be isogenic, due to the horizontal mobility of genetic elements in this species. Therefore, there is a need to determine an isogenic reference for future whole-genome studies. The previous earliest-branching reference isolate, strain MSHR1132, has been described in detail elsewhere [41], and contains an accessory genome which is limited in content. However, this strain is inferior in its capacity as a pathogenic reference in two respects: first, it does not have a propensity for systemic infections; and second, the genetic differences between it and other *S. aureus* strains is 10%, compared to the ~2% genetic difference amongst other sequenced strains of *S. aureus*. Ideally, this genome would be limited in horizontally-transferred genes to those which are absolutely necessary for establishing infections, and would exist within this 2% range.

Herein, we describe one such strain, which we term *S. aureus* Tager 104. Tager 104 was originally isolated at the New Haven Hospital (New Haven, Connecticut) by Morris Tager *et al*. in 1947 from a patient with a cutaneous infection caused by a hemolytic bacteria [35, 37]. As a comparison, the

first strain of methicillin resistant *S. aureus* was isolated in 1961 [38]. Morris Tager and co-workers subsequently proved *S. aureus* induced clotting through expression of secreted factors [132], developed staphylocoagulase purification protocols [36], and initiated preliminary characterization of staphylocoagulase function in contrast to normal physiologic clotting [134, 135]. Currently, staphylocoagulase from *S. aureus* Tager 104 is arguably the most well characterized prothrombin activator studied to date, as it was the original source used to solve the staphylocoagulase structure in complex with both thrombin and its immediate precursor, prethrombin 2 [31]. This recombinant N-terminal fragment of staphylocoagulase from Tager 104 was shown to bind with high affinity (K_D 17-72 pM) to human prothrombin zymogens [136], which was also used to characterize fibrinogen recognition by the prothrombin-staphylocoagulase complex and to determine that two prothrombin-staphylocoagulase complexes bind to a single substrate fibrinogen to mediate cleavage [33]. With the wealth of insight that has been generated from Tager 104, genomic characterization presented here has an interest from both clinical and evolutionary perspectives.

Materials and Methods

Mouse Model of Endocarditis.

We have previously established a model of *S. aureus* endocarditis [62, 137]. For this experiment, we used 12 female C57BL/6 mice obtained from The Jackson Laboratories (Bar Harbor, Maine). Briefly, the model involves anesthetizing mice with 1-3% isoflurane mixed with medical grade oxygen using a vaporizer. Once the mouse is sufficiently anesthetized, as determined by unresponsiveness to tail pinch, the neck and thorax are shaved using electric clippers. Betadine solution is applied to the shaved area and allowed to dry. Animals are given an incision near the carotid artery and the right carotid artery was isolated. The mouse catheter/ suture material (1.4

cm of 4-0 suture sterile thread Ethicon) was inserted down the right carotid artery into the heart and this foreign body was secured in place with a silk suture (8-0 suture thread Ethicon). As the suture material was inserted through the aortic valve, once in position the opening and closing of the valve is apparent by pulsation of the thread. The incisions were closed and animals are allowed to recover for 1 day before bacterial injections, which initiated vegetation development.

Fluorescence Molecular Tomography fused to Computer Tomography (FMT-CT) Detection of Endocarditis.

Between 48-72 hours post infection, we performed FMT-CT imaging to interrogate the magnitude of bacterial involvement in the mice and quantify [Alexa Fluor 680]-Phe-Pro-Arg-Prothrombin (AF680-ProT) deposition within the growing vegetation. FMT imaging was performed on a FMT 2500 system (PerkinElmer) using the 680/700 nm excitation/emission channels, whereby a 3D dataset was reconstructed such that fluorescence/voxel was expressed in nmol/l. To rectify fluorescence signal reporting probe deposition in the heart, anatomic imaging with CT immediately followed FMT. The imaging cartridge containing the anesthetized mouse was placed into the custom-machined Plexiglas holder that supplies isoflurane, warm air, and optimal positioning in the CT (Inveon PET-CT, Siemens, Malvern, Pennsylvania). The CT X-ray source operated at 80 kVp and 500 μA with an exposure time of 370 to 400 ms. The effective 3D resolution was 80 μm isotropic. Isovue-370 CT contrast (Bracco, Princeton, New Jersey) was infused intravenously at 55 μl/min through a tail-vein catheter. The CT reconstruction protocol performed bilinear interpolation, used a Shepp-Logan filter, and scaled pixels to Hounsfield units.

Mouse Model of Systemic Infection.

S. aureus Tager 104 was grown overnight at 37°C in 50 mL of BHI broth in a 125 mL Erlenmeyer flask. Cells were twice washed with 40 mL of filter-sterilized phosphate buffered saline (PBS)

supplemented with 10% glycerol. Centrifuge recovered cells were re-suspended in a final volume of 10 mL of PBS with 10% glycerol. A sample from this inoculum stock was diluted 1:100 in PBS with 10% glycerol in a cuvette for determination of cell density by measuring the absorbance λ_{600nm} using a Beckman Coulter DU800 spectrophotometer. These freshly prepared inoculums were used to inject C57BL/6 mice by tail-vein with 5 x 10^8 CFU. After 24 hours (high dosage), the animals were euthanized.

Histology.

Organs were harvested, stored in 4% paraformaldehyde, and embedded in OCT medium by immersing in 4-methylbutane in a metal canister within a dry-ice bath. Slices (10 microns in width) were made using a Thermo HM550 Cryostat set at -19°C, and fixed to poly-L-lysine coated slides. Slides were then stained using standard antibody, Gram, and H&E staining protocols. Slides were viewed at 100 or 400 times total magnification on a Zeiss Axioskop 40 microscope, and images were taken using a Nikon DS-Fi1 camera head and DS-L3 control unit.

Tager 104 Genome Construction.

Genomic DNA was extracted from Tager 104 using E.Z.N.A. Bacterial D.N.A. kit, and constructed into a bar-coded library using the Nextera DNA sample preparation kit (Illumina). Sequencing was performed using an Illumina MiSeq sequencer for 2 x 150 paired end reads and contigged using CLC Bio v. 4.6.1, as described previously [138]. In order to scaffold these contigs, a sub-library of Tager 104 was constructed for PacBio SMRT sequencing. Two sequencing reactions were performed, and CLC bio contigs were scaffolded using Celera Assembler pipeline on the SMRT analysis 1.3 suite [138].

In order to overcome innate difficulty in genome closure in this initial construction, PacBio reads were instead contigged using SMRT Analysis v. 2.0 Heirarchical Genome Assembly Process

(HGAP) algorithm [139], which produced 8 contigs. In order to close the genome, two Lucigen NxSeq 25kbp mate-pair libraries were constructed and sequenced on an Illumina HiSeq system. PacBio HGAP scaffolds and Lucigen NxSeq paired-end reads were provided to SSPACE [140] to create the final, closed genome. Gap regions in this genome were closed using a combination of the GapFiller algorithm (as part of the SSPACE suite) and Basic Local Alignment Search Tool (BLAST) search against the initial CLC contigs for those which bridge gap regions. This final, closed genome was submitted to the Rapid Annotation using Subsystem Technology (RAST) server [141-143].

To confirm the construction of Tager 104 using an independent method, new libraries were constructed using the Nextera DNA kit and sequenced using Illumina MiSeq 2 x 250 reactions. These results were combined using the SPAdes algorithm [144] for genome closure.

Construction Validation and Analysis.

In order to test the contribution of repeats to the shortcomings in genomic construction, long repeats (> 500 bp) were identified using Nucmer mapping of the Tager 104 genome to itself [145] and selecting for regions with unique locations and proper size. In addition, interspersed repeats and RNA sequences were identified using the RepeatMasker algorithm (www.repeatmasker.org). The coordinates of unique repeats were recorded and provided to Circos version 0.64 (www.circos.ca).

To determine the contribution of Illumina MiSeq contigs (constructed using CLC), PacBio RS sequencing reads, and scaffolds constructed from the combination of the two, results from each assembly were mapped to the Tager 104 genome using Nucmer. The locations of each unique mapping were provided to Circos for visualization.

MLST Analysis.

In order to determine the lineage of Tager 104 in relation to other clinical *S. aureus* strains, we employed multi-locus sequence typing (MLST) analysis [146]. Briefly, the sequences of seven housekeeping genes (Carbamate kinase, *arc;* Shikimate dehydrogenase, *aro*; Glycerol kinase, *glp*; Guanylate kinase, *gmk*; Phosphate acetyltransferase, *pta*; Triosephosphate isomerase, *tpi*; Acetyle coenzyme A acetyltransferase, *yqi*) were determined using *in silico* analysis of the Tager 104 genome with pre-defined primers given by the MLST analysis database, namely arc up - 5' TTG ATT CAC CAG CGC GTA TTG TC -3', arc dn - 5' AGG TAT CTG CTT CAA TCA GCG -3', aro up - 5' ATC GGA AAT CCT ATT TCA CAT TC -3', aro dn - 5' GGT GTT GTA TTA ATA ACG ATA TC -3', glp up - 5' CTA GGA ACT GCA ATC TTA ATC C -3', glp dn - 5' TGG TAA AAT CGC ATG TCC AAT TC -3', gmk up - 5' ATC GTT TTA TCG GGA CCA TC -3', gmk dn - 5' TCA TTA ACT ACA ACG TAA TCG TA -3', pta up - 5' GTT AAA ATC GTA TTA CCT GAA GG -3', pta dn - 5' GAC CCT TTT GTT GAA AAG CTT AA -3', tpi up - 5' TCG TTC ATT CTG AAC GTC GTG AA -3', tpi dn - 5' TTT GCA CCT TCT AAC AAT TGT AC -3', yqi up- 5' CAG CAT ACA GGA CAC CTA TTG GC -3'and yqi dn- 5' CGT TGA GGA ATC GAT ACT GGA AC -3'. These gene fragments were used then to define the allelic assignments for our genome against that of the 155 reference strains used by the site to generate the sequence type. In addition, all ST profiles from the MLST database (mlst.saureus.net) were submitted with Tager 104 to eBurst analysis in order to determine clonal complex assignments.

Proteomic Analysis of Tager 104 Genome.

In order to perform whole-genome comparisons of Tager 104 to reference *S. aureus* strains, a CMG-Biotools virtual machine was graciously provided by Dr. David Ussery [147]. Reference genomes were downloaded from GenBank using the getgbk (i.e. content) function, and converted into FASTA files using saco_convert. These files were then submitted to the prodigalrunner

algorithm to process the open reading frames and create a protein FASTA file. The makebmdest and blastmatrix algorithms were then utilized to construct blast matrices.

Whole-Genome Phylogenetic Analysis.

The annotated Tager 104 genome and the GenBank reference sequence (RefSeq) genomes (Table 4.1) were obtained from RAST and GenBank, respectively. Genomes were submitted to alignment by the Mauve algorithm using default parameters [148]. Orthologs were extracted from the resulting .xmfa file using stripSubsetLCBs (http://gel.ahabs.wisc.edu/mauve/snapshots/2012/2012-06-07/linux-x64) with the minimum LCB length set to 500 bp and the number of genomes set to 32. The result was then converted to FASTA format using the xmfa2fasta.pl script, and the FASTA-formatted alignment was then converted to phylip format using the fasta2phylip.pl script (both parts of the BioPerl package) (www.bioperl.org). In order to construct the phylogenetic tree, a maximum-likelihood tree was constructed using the RAxML program on the CIPRES Science Gateway using default settings (www.phylo.org/portal2/home.action), and the image was constructed using FigTree (tree.bio.ed.ac.uk/software/figtree).

Table 4.1: Bacterial Strains Used In This Study

Strain	Date	Location	Host	Infection	References
71193	2004	United States	Human	Colonization	[149, 150]
04-02981	2004	Germany	Human	Not provided	[151]
11819-97	1997	Denmark	Human	Skin infection	[152]
COL	1960	United Kingdom	Human	Operating theatre	[153]
ECT-R2	2004	Sweden	Human	Not provided	[154]
ED133	1997	France	Ovine	Mastitis	[155, 156]
HO 5096 0412	2005	United Kingdom	Human	Neonatal infection	[157]
JH1	2001	United States	Human	Bloodstream	[158]
JH9	2001	United States	Human	Bloodstream	[158]
JKD6008	2003	New Zealand	Human	Bloodstream	[159]
JKD6159	2004	Australia	Human	Sepsis (IV drug user)	[160]
LGA251	2007	United Kingdom	Avian	Colonization	[161]
M013	2002	Taiwan	Human	Wound	[162]
MRSA252	1997	United Kingdom	Human	Postoperative (septicemia)	[163]
MSSA476	1998	United Kingdom	Human	Osteomyelitis (bacteremia)	[163]
Mu3	1996	Japan	Human	Pneumonia	[164, 165]
Mu50	1997	Japan	Human	Surgical infection	[166]
MW2	1998	United States	Human	Septicemia	[128]
N315	1982	Japan	Human	Pharyngeal smear	[166]
NCTC8325	1960	United Kingdom	Human	Conjunctiva	[167]
Newman	1954	United Kingdom	Human	Osteomyelitis	[39]
RF122	1993	Ireland	Bovine	Mastitis	[168, 169]
S0385	2006	Netherlands	Human	Endocarditis	[170]
T0131	2006	China	Human	Not provided	[171]
Tager 104	1947	United States	Human	Impetigo	This Study
TCH60	2008	United States	Human	Skin sample	http://www.beiresources.org/ProductInformationSheet/tabid/784/Default.aspx?doc=16951.pdf
TW20	2003	United Kingdom	Human	Bacteremia	[40]
USA300_FPR3757	2004	United States	Human	Wrist abscess (IV)	[172, 173]
USA300_TCH1516	2004	United States	Human	Bacteremia	[174, 175]
MSHR1132	2006	Australia	Human	Necrotizing fasciitis	[41]
VC40	2002	Laboratory	N/A	N/A	[176]

MGE-encoded Gene Determinations.

The list of mobile genetic element-encoded virulence factors was determined based on previous studies by our lab and others, and reviewed elsewhere [38]. The RAST server was first searched for the presence of each protein-encoding gene (PEG). In cases where no annotated PEG was detected, the UniProt server (www.uniprot.org) was used to obtain DNA sequences for BLAST analysis. The genetic location of these elements was recorded and provided to CIRCOS for graphic representation.

Genomic islands were determined using the IslandViewer 3 algorithm, available online at http://www.pathogenomics.sfu.ca/islandviewer. Islands were identified based on their genetic content, as reported by the IslandViewer output and BLAST analysis.

For the νSaβ genomic island, genomic profiles were constructed using the RAST server. After determination of the profile, open reading frames were recolored based on their family or function.

Antibiotic Susceptibility Testing.

S. aureus Tager 104 was streaked for a lawn on TSA plates containing 5% Sheep Blood (BD Biosciences). E-test strips for erythromycin (ER), linezolid (LZ), oxacillin (OX), and vancomycin (VA) (bioMérieux, Inc.) were then placed onto the surface of the agar using sterile forceps. All resistance tests were performed under conditions that were naïve for the *S. aureus* Tager 104 stain as to not complicate interpretation by inducing vancomycin-resistance. Bacteria were incubated at 37 C for 17 hours and plates were imaged using an IVIS Lumina XR system. In addition, overnight cultures of *S. aureus* Tager 104 were added at a 1:200 ratio to wells containing brain-heart infusion (BHI) broth and serial diluted concentrations of vancomycin. The 96-well plate was incubated for 17 hours in an incubator and the O.D. $_{600nm}$ was determined using a Thermo Scientific VarioSkan plate reader. Data was subsequently plotted using Graph Pad Prism v4.03.

Results

Tager 104 Forms Potent Endocarditis in vivo

A representative endocarditis mouse imaged by FMT-CT is shown in Figure 4.1 from three vantage points spanning 180 degrees to provide depth of prothrombin recruitment to sites of vegetation development in the mouse. Induction of endocarditis followed injection of *S. aureus* Tager 104. The reporter was an engineered prothrombin analog designed to be catalytically inactivated by a bound inhibitor that also contains a fluorochrome, namely [Alexa Fluor 680]-FPR-prothrombin or [AF680]-ProT [62, 136]. Imaging results support our previous finding that *S. aureus* Tager 104 initiated endocarditis caused the maximal deposition of fluorescent prothrombin probe [62] and that AF680-ProT can be used as a measure of prothrombin activator expression by the pathogen in the resulting vegetation. Gram staining of tissue from a mouse 72-96 hours post infection with *S. aureus* Tager 104 (Figure 4.1D) and immunochemistry using staphylocoagulase polyclonal antibodies (Figure 4.1E) indicate this pathogen is able to express these factors during vegetation formation. Furthermore, our findings show that *S. aureus* Tager 104 has the ability destroy the valve leaflets and that Tager 104 vegetations can grow to a size such that the mass occludes the entire aortic valve.

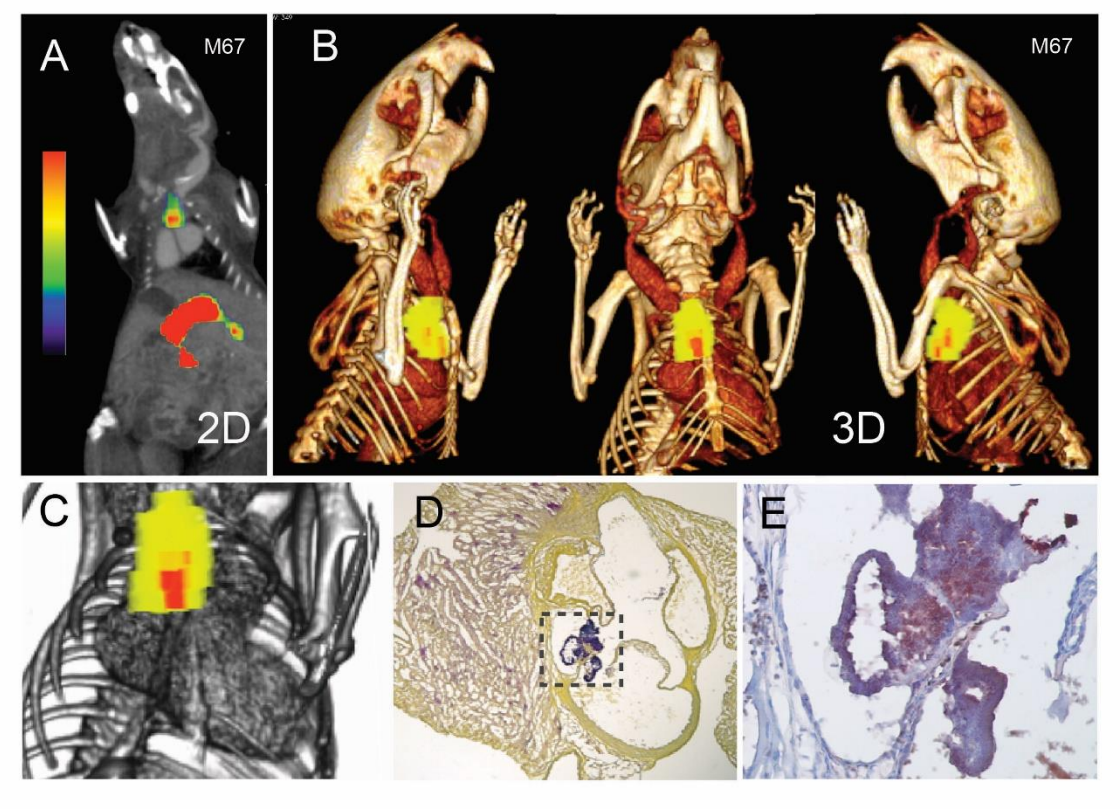

Figure 4.1: *S. aureus* Tager 104 Endocarditis.

A. A 2D alignment of Fluorescence Molecular Tomography fused to Computer Tomography (FMT-CT) images with fluorescence signal emitted from the AF680-ProT is enriched in the upper heart region and liver clearance. *B*. The 3D rendering of the M67 dataset spanning 180° from three vantage points. Localization of the probe is based on the probe binding to prothrombin activators present at the site of the growing *S. aureus* Tager 104 vegetations with vasculature marked by CT contrast. *C*. Isolated view of probe deposition in coronal cross-section of the M67 dataset. Background vascular signal color was eliminated for clarity. *D*. Gram staining for the same mouse showing purple color for bacteria in the aortic valve and myocardium using a 40X total magnification. *E*. Immunohistochemistry for staphylocoagulase the target of the AF680-ProT imaging agent as seen under 400X total magnification. M67 dataset corresponds to 18 h post AF680-ProT injection and a total of 32-36 h post infection.

S. aureus Tager 104 Abscess Formation in Naïve Tissue

Histology indicates that *S. aureus* Tager 104 is able to colonize all tissue with variable coverage at this time point as indicated by Gram staining (Figure 4.2). Bacteria were detected in the alveoli cells of the lung and glomerulus of the kidney. This supports the previous conclusion that Tager 104 represents a systemically-inclined pathogen despite early isolation. Given the ability of Tager 104 to cause both endocarditis and establish abscesses in naive tissue, we sought to better characterize this strain at the whole genome level to ensure that we considering all aspects of virulence determinants prior to side-by-side challenging of hosts with Tager 104 compared to any perceived reference strains.

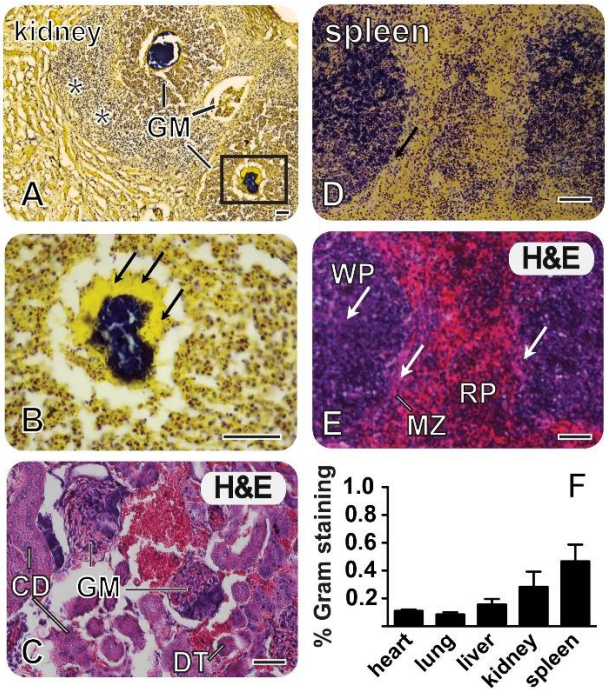

Figure 4.2: *S. aureus* Tager 104 Bacteremia Leading to Multi-organ Septic Foci.

A. Gram staining of C57BL/6 kidney 8 days after injection of 5 x 10^8 CFU *S. aureus* Tager 104. *B.* Higher magnification (400X total) of the indicated by the box in panel A. Arrows indicate the protective fibrin layer that allows the bacteria to thrive in the glomerulus (*GM*) without the threat of clearance by phagocytic immune cells (see area immediately surrounding infected GM shown in Panel A). *C.* Adjacent section of the same kidney stained with Hematoxylin and Eosin (*H&E*); please note the GM is at a slightly earlier stage of the disease. Kidney architecture including the collecting ducts (*CD*), the distal tubule (*DT*), and blood cells (*red*). *D.* Gram staining shows bacteria (*purple*) in all organs examined; splenic abscesses also show thickening of the marginal zone and increased fibrin deposition (arrows). *E.* Adjacent section stained for H&E showing the splenic architecture including the white pulp (*WP*) region, marginal zone (*MZ*) and red pulp (*RP*) region. *F.* Relative organ involvement determined by Gram staining color unmixing using the ImageJ software

Automated Construction of S. aureus Tager 104 Genome.

In order to develop an automated workflow for closure of *Staphylococcus aureus* genomes, Tager 104 genomic DNA was submitted to three separate sequencing technologies: Illumina Miseq 2 x 150 bp reactions, PacBio RS 90-minute movies, and Lucigen NxSeq 25 kbp mate-pair libraries. In all approaches, scaffolds were constructed using one of either Illumina MiSeq or PacBio RS, or a combination of the two (hybrid assembly).

We first sought to close the hybrid assembly scaffolds previously produced for Tager 104 [138]. Therefore, we submitted these scaffolds to the SSPACE algorithm with the Lucigen NxSeq mate pairs. This approach successfully closed the genome, producing the 2.8 Mbp *S. aureus* genome (data not shown). However, this approach left large gaps in the assembly, which require significant manual closure.

Therefore, we instead tested the ability of PacBio RS sequencing reactions in combination with NxSeq mate-pair libraries. Scaffolds were produced using the newer Hierarchical Genome Assembly Process (HGAP) assembly algorithm on the PacBio SMRT Analysis suite (V. 2.0) [139]. This algorithm produced 8 scaffolds from raw sequencing data, with an N50 value of 1,028,373 bp. These scaffolds were then closed using deduplicated NxSeq reads using SSPACE. Gaps resulting in the construction were then filled by aligning CLC-constructed Illumina contigs to the genome using Basic Local Alignment Search Tool (BLAST) search. The result was a closed 2.82 Mbp genome (Figure 4.3A).

To test the robustness of Illumina MiSeq reactions in combination with NxSeq mate-pair libraries, Tager 104 genomic DNA was sequenced using Illumina 2 x 250 sequencing reactions, and resulting contigs were scaffolded using Lucigen NxSeq reactions. The resulting closed genome was identical to that produced by the HGAP-SSPACE assembly pipeline (Figure 4.3B).

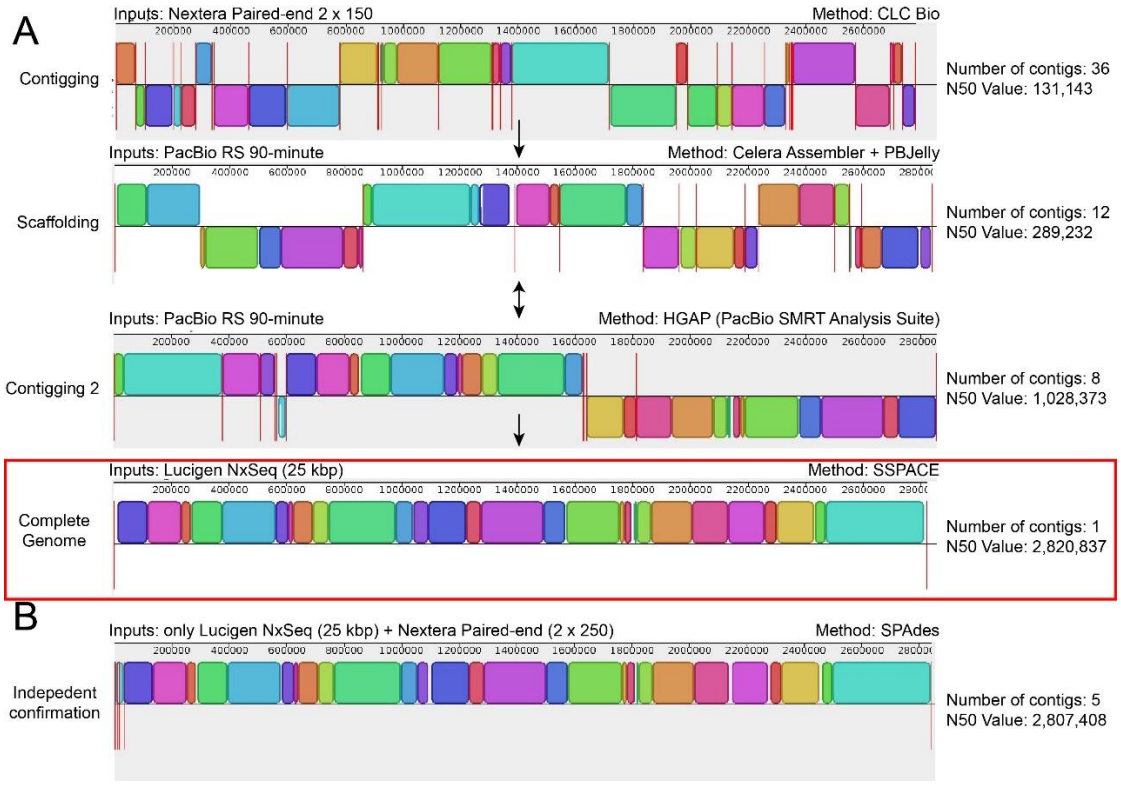

Figure 4.3: Construction of the Tager 104 Genome

A. The Tager 104 genome was constructed using indicated methods, with each step incorporating new genomic data (*Inputs*). First, Tager 104 paired-end reads from Illumina MiSeq sequencing were submitted to CLC Bio (*Contigging*). Resulting contigs were scaffolded using reads from PacBio RS sequencing (*Scaffolding*). Due to the inability to further close scaffolds, PacBio RS data alone was instead contigged using HGAP assembly (*Contigging 2*). The eight remaining contigs were then closed using Lucigen NxSeq mate-pair libraries with the SSPACE algorithm. Single-headed arrows represent events where resulting data from the previous step was submitted to alogrithms in the subsequent step (*Method*). Steps in which results were interchangeable are indicated by two-headed arrows. The assembly is described in more detail in "Materials and Methods." *B.* In order to confirm the completed genome, Lucigen NxSeq libraries were submitted with separate 2 x 250 paired-end reads in an independent SPAdes construction devoid of previous read data from panel *A*. Results were identical to the finished genome.

Repeat Regions Created Difficulty in Automated Closure

In order to determine the shortcomings of individual methods on the closing of the Tager 104 genome, as a model for *S. aureus* genomes, each contig or scaffold was mapped to the final genome using Nucmer and visualized using Circos (Figure 4.4). In addition, repeats were determined by mapping the Tager genome to itself and selecting for matches with unique locations and a size > 500 bp. Interspersed repeats and RNA sequences were determined using the RepeatMasker algorithm. Based on the analysis, Illumina MiSeq 2 x 150 reads did not properly bridge areas of the genome rich in repeats. More specifically, repeats > 500 bp (Figure 4.4, Region 1) and RNA repeats (Figure 4.4, Region 2) caused the greatest issues in construction, as visualized. In addition, PacBio RS coverage was low across these areas.

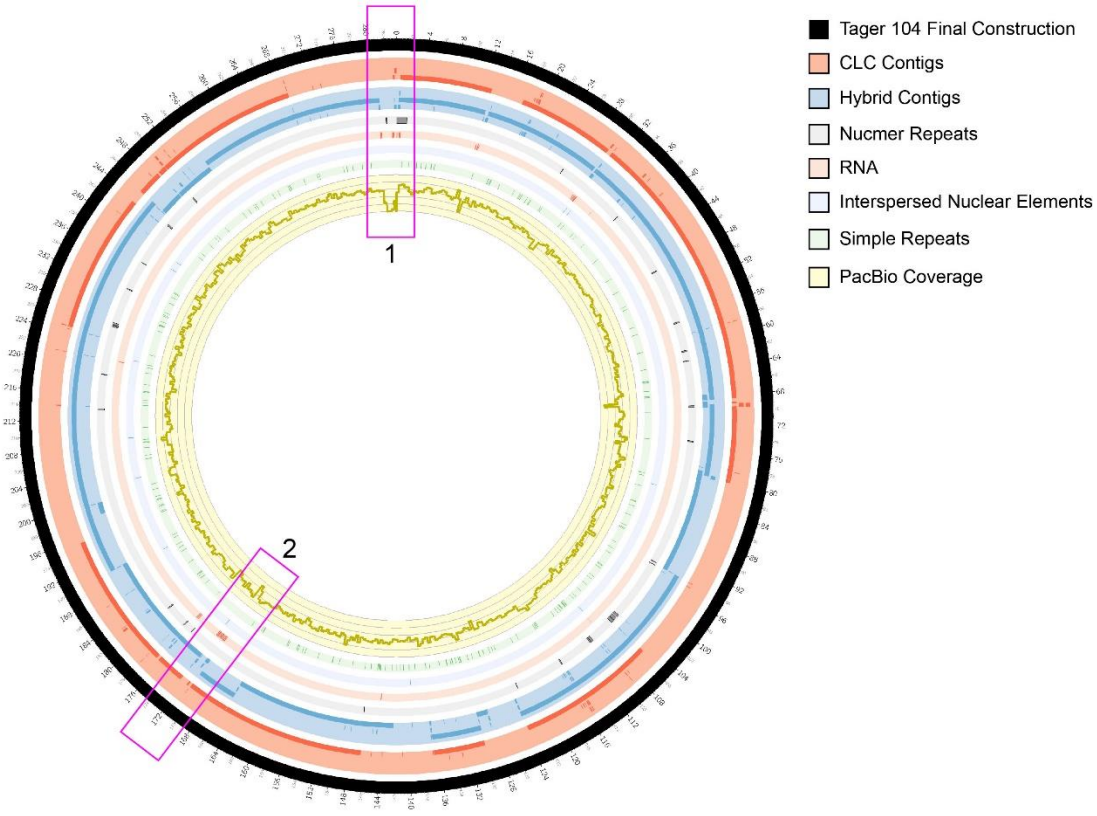

Figure 4.4: Graphical Depiction of Tager 104 Assembly

The Tager 104 genome was scaffolded using three separate technologies: contigs from MiSeq reactions (*red tiling*), PacBio RS reads (*yellow histogram*), and scaffolds produced from the hybrid assembly of both (*blue tiling*). However, assembly of the Tager 104 genome using these technologies were found insufficient alone, due to the presence of repeat regions. More specifically, long repeats (> 500 bp) create errors in de Bruijn graph untangling (*gray tiling*). In addition, the exact placement of short repeats, such as intersperced nuclear elements (*light blue tiling*), RNA sequences (*light red tiling*) and simple repeats (*light green tiling*) create errors in construction. Areas rich in repeats and low in sequencing coverage (*Regions 1 and 2*) reveal difficulties in genomic construction. These difficulties were overcome by providing the construction algorithms with 25 kbp mate-pair information, provided by Lucigen NxSeq libraries.

Proteomic Analysis of S. aureus Reference Sequences

Delineation of the *S. aureus* strains is typically achieved using the multi-locus sequence typing analysis. Briefly, the analysis utilizes sequence heterogeneity among seven housekeeping genes in order to provide a higher resolution phylogenetic analysis. Reference genomes were obtained from the GenBank RefSeq database (Table 4.1). Multi-locus sequence and *spa* typing of *S. aureus* Tager 104 revealed it to be ST49-t208, the predicted founder of clonal complex 49 (CC49). CC49 was also shown to contain ST138, ST1693, ST1937, and ST2273 based on eBurst analysis (data not shown) [177].

In order to determine the similarities of protein content in completed sequences of *S. aureus*, each genome was called for open reading frames using the prodigalrunner algorithm and compared using the blastmatrix algorithm on CMG-Biotools version 2.2. The results (Figure 4.5) indicate the Tager 104 proteome is closest related to M013, an ST59, *pvl*-positive MRSA strain isolated from a wound in Taiwan [162]. Interestingly, the greatest proteomic differences were seen for strain S0385, a typically ovine-associated strain isolated from a human case of endocarditis [170]. In addition, the proteomic similarities for organisms in the same multi-locus sequence type (MLST) was shown to range from as low as 84.2% in ST5 and as high as 99.9% in ST105. This high end is to be expected, as JH1 and JH9 were sequentially isolated from the same patient undergoing vancomycin therapy [178]. Despite differences in sequence type, S. aureus 04-02981 shows > 94% similarity to ECT-R 2, Mu3, Mu50, N315, JH1, and JH9, supporting previous discoveries [151].

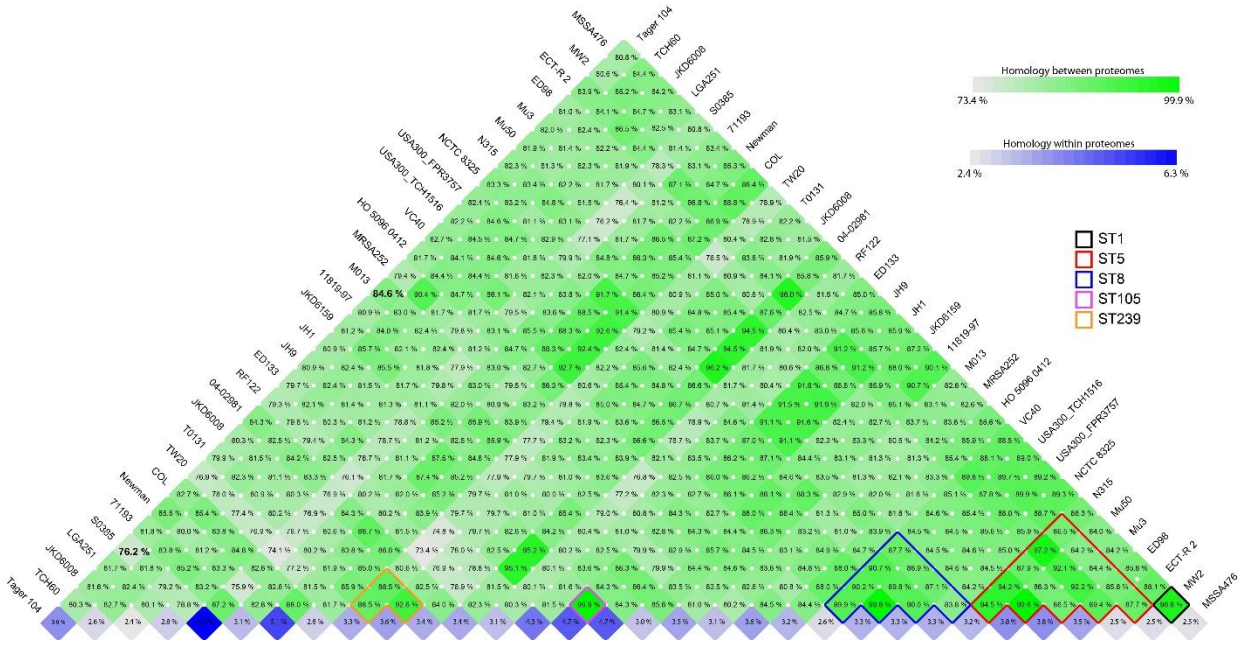

Figure 4.5: Proteomic Analysis of *S. aureus* Reference Sequences

S. aureus reference sequence GenBank files were downloaded from the NCBI server and genes were called using the prodigalrunner algorithm on a CMG-Biotools system. Protein similarity between two genomes is indicated (*green box*), as well as the protein similarity within a single genome (*blue box*). Genomes were organized based on sequence type (ST) analysis, and those from the same ST class are indicated by colored rectangles.

Tager 104 is an early-branching ancestor to modern clinical strains.

To determine the lineage of S. aureus reference strains in relation to Tager 104, genomes were submitted to Mauve alignment and conserved segments were stripped and concatenated for each genome for phylogenetic analysis by RAxML. This alignment and phylogeny is useful in that it only utilizes genetic regions detected in all reference strains, essentially ignoring elements of the accessory genome unique to individual strains. This analysis (Figure 4.6A) determined Tager 104 to be an early-branching strain of the *S. aureus* lineage, and indicated a low genetic divergence to other reference strains. In order to determine the regional evolution of *S. aureus*, this analysis was then repeated for strains isolated in the United States (Figure 4.6B). Tager was shown to be at the root of this analysis.

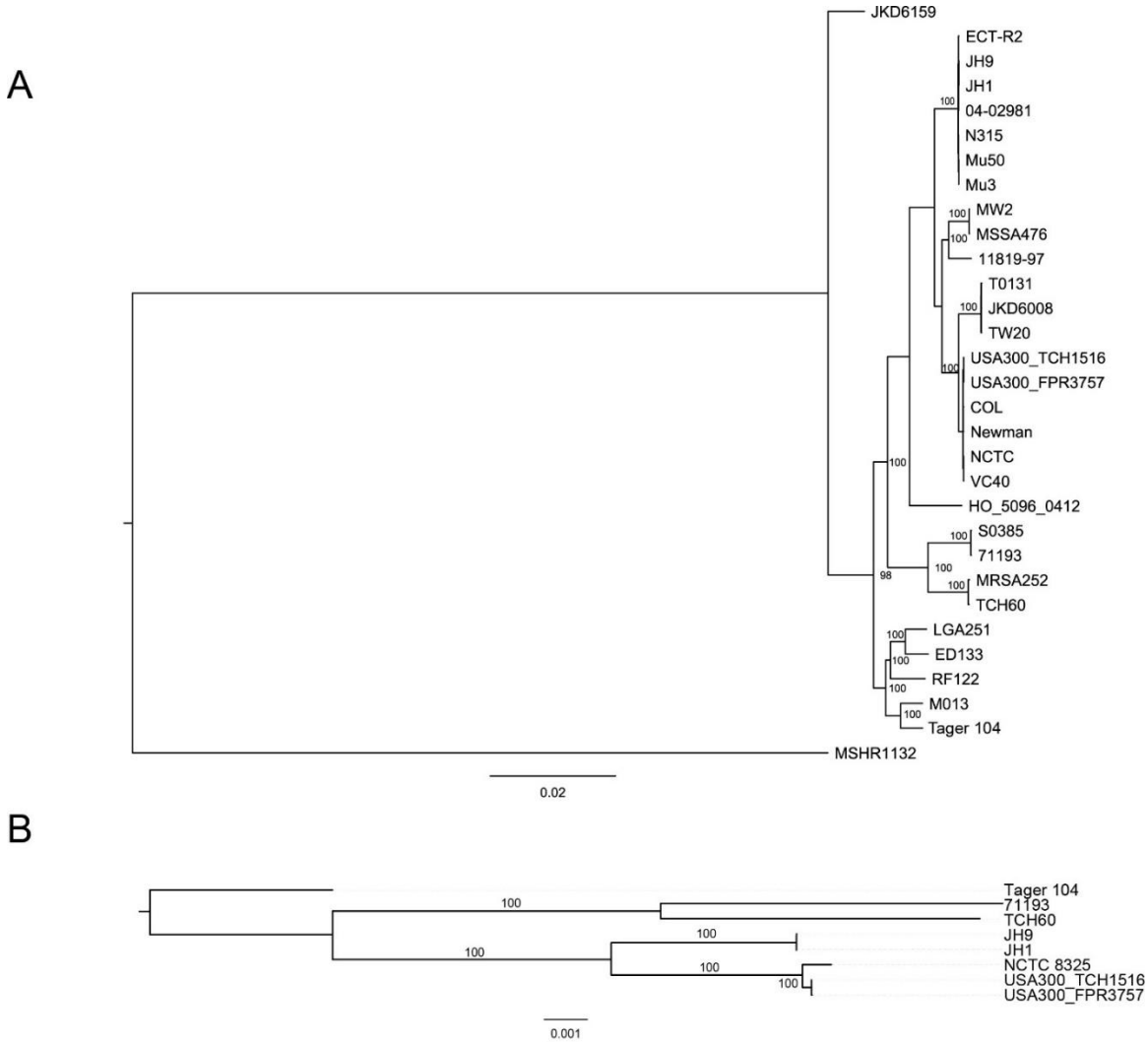

Figure 4.6: Whole-Genome Phylogenetic Analysis of *S. aureus* Reference Strains

Whole-genome phylogenetic analysis of all *S. aureus* reference strains (*A*) or isolates from the United States (*B*) based on Mauve alignment and RAxML analysis. Bootstrap values represent the result of 100 trials.

Tager 104 Contains a Limited Set of MGE-encoded Virulence Factors Targeted at Host Immunity

As Tager 104 was able to survive in septic murine models, and due to its early lineage in *S. aureus* evolution, we investigated the genomic content of mobile genetic element-encoded virulence factors, which have been previously identified and characterized [38]. The results (Figure 4.7) demonstrate a limited set of factors focused primarily on survival during bacteremia.

Tager 104 contains νSaα-encoded Staphylococcal superantigen-like SSL, as well as νSaβ-encoded leukotoxin D and E (*lukD* and *lukE*, respectively). These factors are known to inhibit elements of innate immunity [22]. The νSaβ island was also shown to contain the hyaluronate lyase gene (hysA), which degrades hyaluronic acid found in host extracellular matrix.

Genetic analysis also indicated early exposure to phage-carried virulence factors. Tager 104 contained the Staphylococcal inhibitor of complement (*scn*), which targets innate immunity by inhibiting phagocytosis by neutrophils. Tager 104 was shown to contain the *hlb*-interrupting phage described in the genomes of modern clinical strains, such as N315, Mu50, MW2, NCTC8325, MSSA476, MRSA252, USA300, JH1, JH9, and Newman [179], and was defined here as φTGR1. Interestingly, Tager 104 also contained the Panton-Valentine leucocidin (*lukF-PV* and *lukS-PV*), a phage-transferred pore-forming leukocyte toxin linked to necrotic infections. Tager 104 only contained one enterotoxin, determined by BLAST analysis as an exact match to the phage-transferred enterotoxin P [38].

In addition, Tager 104 was determined to contain one unknown prophage and four unknown genomic islands, indicated as φTGR2 and TGR1 through TGR4, respectively. The contents of these islands are summarized in Table 4.2 and Table 4.3, and were limited to hypothetical and/or phage proteins.

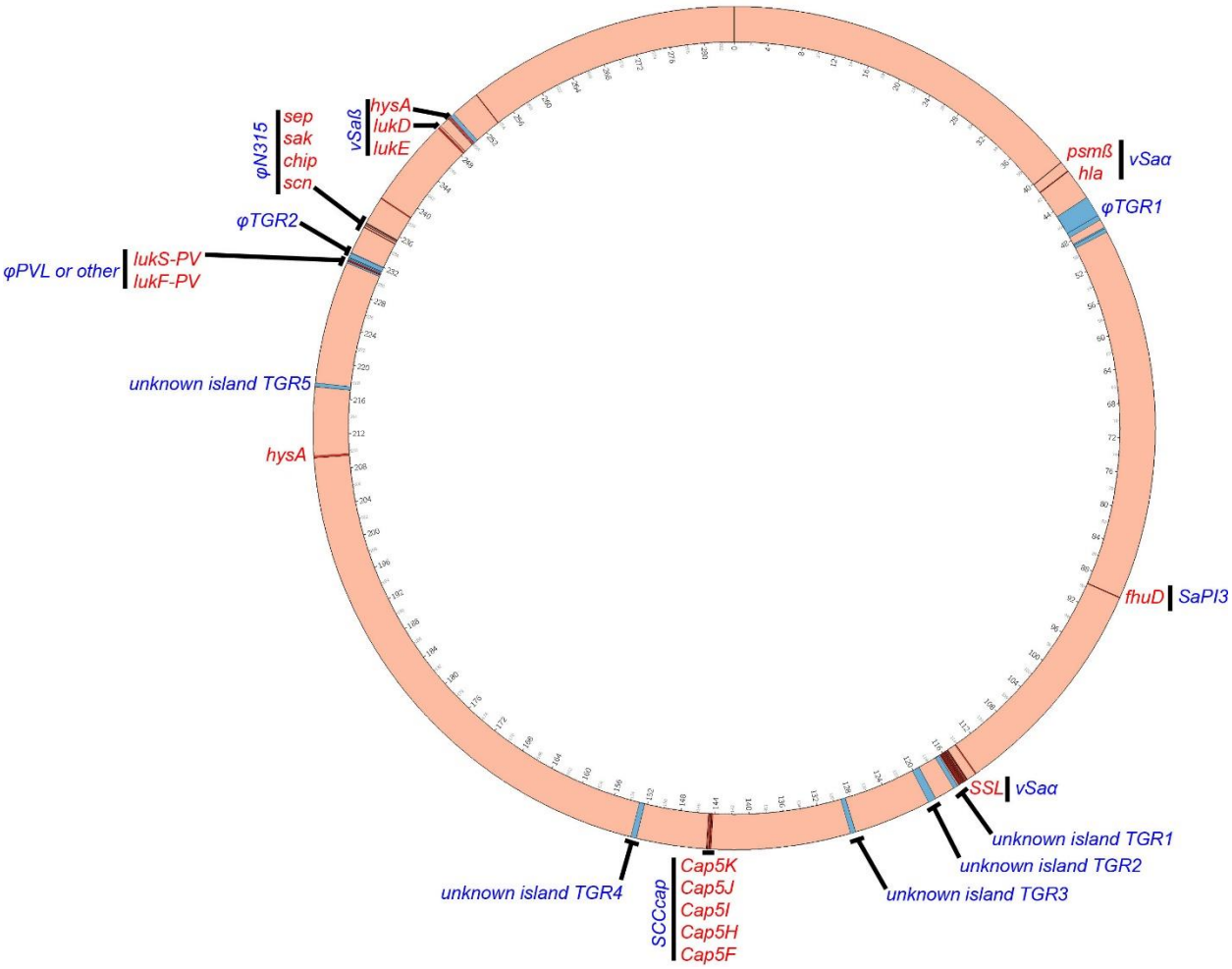

Figure 4.7: Mobile Genetic Element-Encoded Virulence in *S. aureus* Tager 104.

Mobile genetic element (MGE)-encoded genes relevant for virulence in S. aureus were selected based on prior knowledge of S. aureus genomics. Those elements detected via BLAST analysis in the Tager 104 genome are indicated by locus (*red bands*) as well as gene name (*red labels*), and the genomic island responsible for carrying these elements are grouped (*black bar*) and indicated beside the gene name (*blue label*). In addition, islands were detected using the IslandViewer 3 software, and are indicated by locus (*blue bands*) and given identifiers (*blue label*) based on gene identities and BLAST result.

Table 4.2: Genomic Island Contents in S. aureus Tager 104

Genomic Location	Name	BLAST Result	Gene Annotations
1,165,336 - 1,170,248	TGR 1	SA268, SA40, SA957, M013	ORF031
			ORF021
			Phage lysin, N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
			Phage holin
			Phage protein
			Tail fiber protein [SA bacteriophages 11, Mu50B]
			Phage lysin, N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
			Phage protein
			Hypothetical protein, phi-ETA orf58 homolog [SA bacteriophages 11, Mu50B]
			Phage protein
			FIG01108548: hypothetical protein
			Putative major teichoic acid biosynthesis protein C
			Phage minor structural protein
			phi 11 orf43 homolog [SA bacteriophages 11, Mu50B]
			Phage tail length tape-measure protein
			phi 11 orf41 homolog [SA bacteriophages 11, Mu50B]
			Phage protein
			Phage tail protein
			Phage protein
			Phage protein
			Phage protein
			Phage transcriptional terminator
			Phage major capsid protein
			Phage capsid and scaffold
1,192,427 - 1,201,074	TGR 2	6850	FIG01108790: hypothetical protein
			FIG01108566: hypothetical protein
			FIG01107981: hypothetical protein
			Integrase, superantigen-encoding pathogenicity islands SaPI
			FIG01108853: hypothetical protein
			FIG01108840: hypothetical protein
			hypothetical protein
			hypothetical protein
			hypothetical protein

				hypothetical protein
				hypothetical protein
				hypothetical protein
				Putative terminase, superantigen-encoding pathogenicity islands SaPI
				Hypothetical SAV0799 homolog in superantigen-encoding pathogenicity islands SaPI
1,285,693	-	TGR 3	SA40, SA957, M013	FIG01108751: hypothetical protein
1,291,359				FIG01108312: hypothetical protein
				FIG01108876: hypothetical protein
				FIG01107881: hypothetical protein
				Lmo0069 homolog within ESAT-6 gene cluster
				Lmo0069 homolog within ESAT-6 gene cluster
				FIG01108656: hypothetical protein
				FIG01108452: hypothetical protein
				FtsK/SpoIIIE family protein, putative EssC component of Type VII secretion system
1,523,588	-	TGR 4	SA957	acetyltransferase (GNAT) family protein
1,529,604				HTH-type transcriptional regulator LrpC
				hypothetical protein
				hypothetical protein
				Bipolar DNA helicase HerA
				FIG036446: hypothetical protein
				hypothetical protein
2,170,899	-	TGR 5	FORC_001, MRSA252	FIG01107943: hypothetical protein
2,174,905				FIG01107943: hypothetical protein
				FIG01107943: hypothetical protein
				Transcriptional regulator, Cro/CI family protein transposon-related
				FIG01107943: hypothetical protein
				FIG01107943: hypothetical protein
				FIG01107943: hypothetical protein

Table 4.3: Prophage Content in S. aureus Tager 104

Genomic Location	Name	BLAST Result	Gene Annotations
447,213 - 490,078	ϕTGR1	ϕNM1-4	ORF031
			ORF021
			Phage lysin, N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
			Phage holin
			Phage protein
			Tail fiber protein [SA bacteriophages 11, Mu50B]
			Phage lysin, N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)
			Phage protein
			Hypothetical protein, phi-ETA orf58 homolog [SA bacteriophages 11, Mu50B]
			Phage protein
			FIG01108548: hypothetical protein
			Putative major teichoic acid biosynthesis protein C
			Phage minor structural protein
			phi 11 orf43 homolog [SA bacteriophages 11, Mu50B]
			Phage tail length tape-measure protein
			phi 11 orf41 homolog [SA bacteriophages 11, Mu50B]
			Phage protein
			Phage tail protein
			Phage protein
			Phage protein
			Phage protein
			Phage transcriptional terminator
			Phage major capsid protein
			Phage capsid and scaffold
			Phage protein
			Phage protein
			Phage portal protein
			Phage terminase, large subunit
			Phage terminase, small subunit
			Integrase regulator RinA
			Phage protein
			Transcriptional activator rinB, phage associated
			Hypothetical protein, SAV0877 homolog [SA bacteriophages 11, Mu50B]

			hypothetical protein
			ORF058
			Phage protein
			ORF077
			Phage antirepressor protein
			Cro-like repressor [SA bacteriophages 11, Mu50B]
			Phage repressor
			hypothetical protein within prophage
			Hypothetical protein, SAV0849 homolog [SA bacteriophages 11, Mu50B]
			Phage excisionase
			Phage DNA invertase
2,308,294 2,313,700	-	ϕPVL or other	Acetylornithine deacetylase (EC 3.5.1.16)
			Leukocidin LukS-PV
			Leukocidin LukF-PV
			Beta-hemolysin
			Phage integrase
2,313,771 2,320,128	-	ϕTGR2 ϕNM3	Phage integrase
			Phage protein
			glycosyl transferase
			hypothetical protein within prophage
			Phage protein
			DNA helicase, phage-associated
			Phage repressor
			DNA-binding protein, phage associated
			hypothetical protein within prophage
			hypothetical protein within prophage
			Phage antirepressor protein

Tager 104 Shows Intermediate Clinical Adaptations

One group of pathogenicity islands that are non-phage and non-SCC genomic islands are termed νSa islands, and typically contain a combination of virulence factors and either an intact or remnant recombinase [39]. Two such islands, termed νSAα and νSAβ, were previously found to be allelic, and are therefore typed and used in conjunction with sequence type analysis for determination of strain radiation. Comparison between the profiles of the νSAβ genomic island for one community-acquired MRSA (CA-MRSA) strain (COL), three hospital-acquired MRSA (HA-MRSA) strains (MRSA252, Mu3, JH1), and two ruminant host MSSA strains (S0385, RF122), as well as Tager 104 (Figure 4.8) demonstrated *S. aureus* Tager 104 has lost the majority of the *bsa* locus with the exception of *bsaG*. However, *S. aureus* Tager 104 has not acquired enterotoxin genes seen in other strains adapted for the hospital setting (HA-MRSA). The loss of this *bsa* locus, coupled with the gain of enterotoxins has been previously hypothesized to indicate a shift from the environmental niche to one of hospital acquisition [180]. Therefore, *S. aureus* Tager 104 may represent an intermediate strain for the selection of this νSAβ profile in the hospital setting.

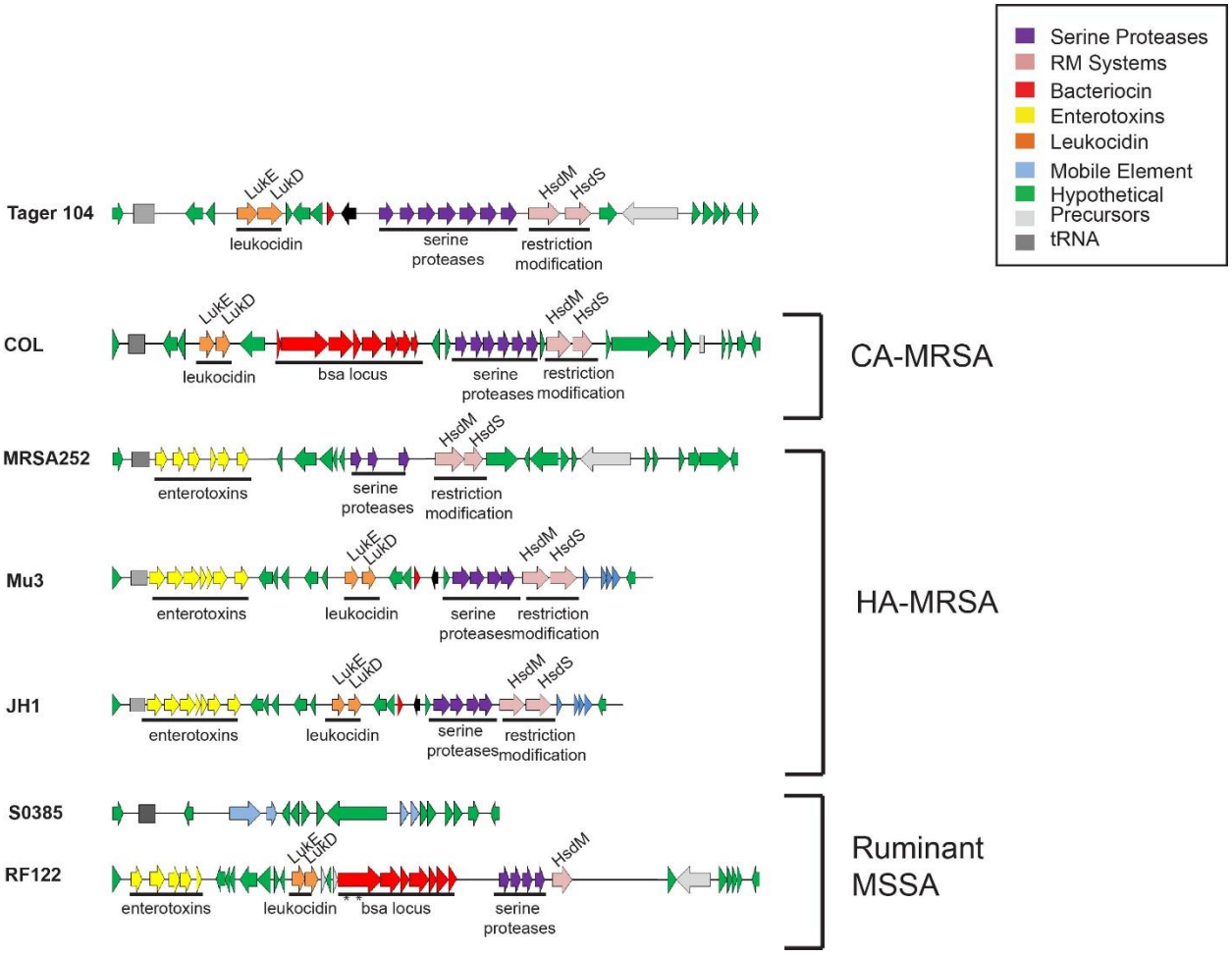

Figure 4.8: Tager 104 vSaβ Genomic Island Shows Intermediate Clinical Adaptation.

Representatives from each subtype of the vSaβ locus were selected and compared using the RAST server. Current strains of community acquired MRSA (*CA-MRSA*) contain the *bsa* locus (*red arrows*), which encodes a bacterial antibiotic. In contrast, hospital-acquired MRSA (*HA-MRSA*) have lost the *bsa* locus and instead gained enterotoxins (*yellow arrows*), an adaptation which has been previously associated with a selection for hospital settings. Results indicate Tager 104 has lost the *bsa* locus, but lacks the enterotoxin cluster (yellow) that has also been associated with this shift. In contrast, ruminant-host MSSA shows a wide range of these genotypes.

S. aureus Tager 104 Resistance Profile

In order to determine the potential resistance in Tager 104, known MGE-transferred resistance cassettes [38] were searched against the closed genome using BLAST. No resistance elements were detected (data not shown).

Minimum inhibitory concentrations (MIC) for *S. aureus* Tager 104 were determined by challenging with E-test strips for linezolid, erythromycin, oxacillin and vancomycin. MICs for antibiotics determined by E-test strip analysis are shown in Fig. 9; *S. aureus* Tager 104 demonstrated resistance to linezolid at 1 µg/mL, erythromycin at 0.64 µg/mL, oxacillin at 0.19 µg/mL and vancomycin at 2 µg/mL. The vancomycin-intermediate susceptibility phenotype of *S. aureus* Tager 104 was confirmed by solution-based test in a 96-well format; results confirm the MIC to be about 3 µg/mL. *S. aureus* Tager 104 colonies on brain heart agar plates are well-formed with the absence of satellite projections and an overall white color, which is unlike other vancomycin-intermediate strains described previously [159].

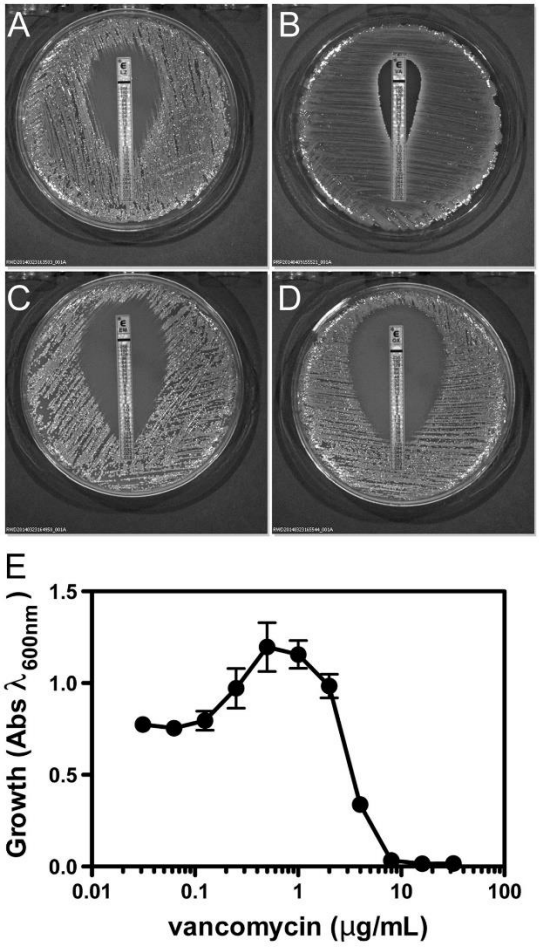

Figure 4.9: *S. aureus* Tager 104 Susceptibility to Antibiotic Therapy Demonstrates Predation of Resistance Development.

A-D. Antibiotic susceptibility testing on sheep blood agar (SBA) plates demonstrated *S. aureus* Tager 104 is susceptible to erythromycin (*A*), linezolid (*B*), oxacillin (*C*), and vancomycin (*D*) when challenged with E-test strips. Minimum Inhibitory concentrations shown are: erythromycin, 64 µg/mL; linezolid, 1 mg/mL; oxacillin, 19 µg/mL; and vancomycin, 2-3 mg/mL. *E.* The effects of vancomycin concentrations on Tager 104 growth were confirmed in solution in 96-well plates and indicate a vancomycin intermediate susceptibility phenotype. This experiment is a combination of 8 replicates and was performed as described in Material and Methods.

Discussion

S. aureus infections range in presentation from localized skin infections to more life-threatening osteomyelitis, endocarditis and sepsis. With the advent of next-generation sequencing, analysis is now possible for individual genomes across the spectrum of infections and has revealed that certain *S. aureus* exhibit considerable changes in their gene make-up thus reflecting selective pressures to infect certain host preferentially [168]. The acquisition of mobile-genetic elements and smaller changes in genes known as nucleotide polymorphisms together confer the apparent selective advantages needed for these strains to survival in a given host. Directly comparison of sequence types (ST) Horizontal gene transfers in *S. aureus* strains fuels this adaptation. For example, *S. aureus* Newman, isolated in 1952, has been shown to contain four prophages (φNM1 – φNM4) and one additional νSa island, νSa4 [39]. Lack of these prophages was associated with decreased abscess formation in organs. However, strains such as MW2 and N315 lack these prophages [39]. Therefore, future characterization of essential elements of *S. aureus* systemic virulence requires an isogenic strain, which is limited in genomic island content and has a propensity for systemic infection.

Our previous results [30], as well as those reported here, indicate that *S. aureus* Tager 104 is a potent strain for the formation of bacterial endocarditis, a serious complication of bacteremia (Figure 4.1). Indeed, Tager 104 was originally isolated and maintained for decades, due in no small part for its potent ability to form clots in human, rabbit, and sheep blood [112]. This clotting is now known to be caused by a coordinated effort of two non-proteolytic activators of prothrombin, namely staphylocoagulase [31] and von Willebrand factor binding protein (vWbp) [32] [34]. Both of these factors activate prothrombin by forming a reversible complex that cleaves their substrate fibrinogen to provide the fibrin barrier, similar to those seen in the focal abscesses seen here in multiple organs, including the kidney and spleen (Figure 4.2) that were formed by

simple induction of diffusive bacteremia. Based on these observations, combined with the early date of its isolation, we sought to determine the ancestry and isogenicity of Tager 104 for future evolutionary and correlation studies against other *S. aureus* strains.

Our goal in sequencing was to find the simplest and most inexpensive combination of technologies to automate the closure of the Tager 104 genome, as a reference for future *S. aureus* genomic sequencing efforts. We began with Illumina MiSeq 2 x 150 paired-end libraries, constructed using the Nextera DNA kits. In order to close the resulting contigs, literature pointed to the use of PacBio RS technology to bridge across complex regions using longer reads [51]. PacBio reads were corrected using Illumina MiSeq paired-end reads, and scaffolded to create the final construction. The result from this "hybrid assembly" was twelve scaffolds, which have been previously described [138]. However, no combination of algorithms or manual assembly were able to close the genome. New advancements in PacBio construction, ushered by the release of the newer HGAP assembly algorithm, led us instead to reconstruct the genome using raw PacBio data alone, producing 8 scaffolds, which similarly would not reduce, and bridging them manually proved problematic. We therefore decided to approach the closing of the genome using greater distances between the paired reads, in order to provide additional information to the construction algorithms. Two Lucigen 25 kbp mate-pair libraries (produced using two separate digestion regiments) were sufficient to close the Tager 104 genome from this HGAP-assembled data. Furthermore, Illumina 2 x 250 paired-end libraries were constructed, sequenced, and contigged. Upon addition of the Lucigen 25 kbp mate-pair libraries, the final genome was produced, and was identical to that of the PacBio technology. Therefore, these mate-pair libraries in combination with the free-ware SPAdes software provide a streamlined method for genomes assembles and specific closure of *S. aureus* genomes.

Ideally, for Tager 104 to serve as an isogenic reference for *S. aureus* systemic infections, it will need to have: (1) be a clinical sample, (2) be early branching member on the *S. aureus* phylogenetic tree, and (3) be limited in MGE content to those which are necessary for systemic survival (*i.e.*, immune system evasion). To address the first, Tager 104 has been maintained from primary freezer stocks since its isolation in the late 1940's. This isolation is the earliest in the literature of *S. aureus* genomes to date. Tager 104 shows ancestral phylogeny similar to other methicillin-resistant and methicillin-susceptible strains that have been characterized in the literature [39, 152, 155, 159, 160, 162, 170, 171, 173]. Our results (Figure 4.6) indicate Tager 104 as an early-branching member of the *S. aureus* phylogeny, in agreement with its early isolation date. The previous early-branching strain, MSHR1132, shows a 10% genomic divergence from other *S. aureus* strains, which themselves show an average 2% divergence. *S. aureus* Tager 104 was determined by average nucleotide identity (ANI, http://enve-omics.ce.gatech.edu/ani/) to match this 2% divergence from other clinical strains.

Lastly, Tager 104 shows a limited set of MGE content primarily focused on the evasion of host innate immunity. Tager 104 was shown to contain the *hlb*-interrupting phage described above, labeled as φTGR1, as well as two additional phages. One of these phages contained the Panton-Valentine leukocidin, indicating the earliest acquisition of this factor in the literature. The other phage was shown to carry no additional genes aside from those necessary for phage construction and replication (Table 4.2). Tager 104 was shown to contain the νSaα and νSaβ islands, found in all *S. aureus* genomes [39]. Although Tager 104 was also detected to contain four additional potential genomic islands, their contents were hypothetical proteins. Therefore, the virulence-associated genomic island composition of Tager 104 is limited to those which are found in all

genomes (νSaα, νSaβ), the hlb-interrupting phage found in most all pathogenic *S. aureus* (φTGR1), and the Panton-Valentine leucocidin-carrying phage.

Our *S. aureus* Tager 104 genome also indicated that Tager 104 pre-dates the acquisition of the staphylococcal cassette chromosome *mec* (SCC*mec*) cassette and may serve to bridge certain clinical methicillin-resistant *S. aureus* (MRSA) and methicillin-susceptible *S. aureus* (MSSA) strains (Fig. 6). Supportive to this hypothesis is the recent observation of an ST49-t208 strain that has become MRSA [181], potentially through the acquisition of the *mecALGA251* cassette [182]. In concordance with this observation, the Tager 104 clade split from the common ancestor before the development of clades containing COL, an early-branching MRSA strain [153], and those containing JH1 and JH9, which were previously described as strains used for the determination of mechanisms for the acquisition of vancomycin resistance in a patient with endocarditis [178]. These results indicate Tager 104 represents an intermediate departure in the evolution of multi-drug resistant *S. aureus*. In addition, the results of the phylogenetic analysis indicate that JKD6159 represents a methicillin-resistant strain, which radiated from this lineage at a point before Tager 104. Antibiotic susceptibility of *S. aureus* Tager 104 presented here using clinically approved E-test analysis corroborated genomic data (Fig. 4.9) and indicated Tager 104 displayed a borderline vancomycin intermediate *S. aureus* phenotype (Fig. 4.9B).

It has been recently suggested that the acquisition of these mobile genetic elements may be driven by "glycocodes" encoded by the teichoic acid structure of *S. aureus*, as these structures are recognized by bacteriophages during transduction events [183]. With this in mind, we investigated the locus of teichoic acid biosynthesis in *S. aureus* Tager 104, COL (as a representative of early MRSA strains) and JH1 (as a representative of early VRSA strains) with the hypothesis that homology in these strains indicate a potential phage-mediated acquisition of mobile genetic

elements. The results indicate that the genomic makeup of this biosynthesis cluster is identical in these three strains. Therefore, the predating ancestor to the division of *S. aureus* Tager 104, or Tager 104 itself, has the potential to have acquired resistance elements from bacteriophages leading to the evolution of more recent clinical MRSA and VRSA strains, in agreement with this phylogenetic analysis (Fig. 4.6). As expected, the PS187 genome described previously showed a very distinct profile, furthering the previous conclusion that this strain is distantly related to other known *S. aureus* lineages. Further analysis will be required to investigate the origin of resistance transfer into Tager 104 and its contribution to *S. aureus* virulence.

Identification of a sufficient reference for systemic pathogenesis by *S. aureus* and acquisition of multi-drug resistance are crucial to give context and scope to the small nuclear polymorphisms (SNPs) granting adaptive advantages to the pathogen *in vivo*. Whole-genome association studies (WGAS) can then associate these SNPs with phenotypic information, such as that provided by our imaging modalities herein, to determine statistical validity. This process has been utilized previously to determine SNPs which grant an intermediate resistance to vancomycin [54, 133]. Within, we have demonstrated that Tager 104 displays a virulent phenotype during infections *in vivo*. Phylogenetic analysis, as well as a set of mobile genetic element-encoded factors limited to that which is necessary for systemic survival, suggest Tager 104 is a strong reference for the future studies in *S. aureus* systemic virulence.

Once these future studies have identified and characterized the mutations associated with disease progression using this WGAS approach, molecular imaging can be utilized to determine the exact timing and location of these factors, such as that seen in Chapter 2. These future comparisons will rule out redundant factors, and focus on the targeting of essential virulence factors through novel anti-virulence factor approaches.

5. Conclusions, Discussion, and Future Directions

In this dissertation, we have investigated the systems biology of invasive infections in two approaches: first, in the development and characterization of genetically-engineered models for *S. pyogenes* dissemination and light production; and second, in the characterization of a reference strain of *S. aureus* potent for bacteremia and infective endocarditis. In addition, we have demonstrated the combinatorial use of molecular imaging and next-generation sequencing to determine the necessary virulence factors in invasive infections. In the case of *S. pyogenes*, molecular imaging data from Chapter 2 and Chapter 3 were correlated with previous metabolic models in order to determine the means by which cellular physiology affects light production. The impact of this project is two-fold: first, it outlines the necessity of determining the physiological background of pathogens before cloning the *luxABCDE* cassette into the strain for in vivo study; and second, it demonstrates, for the first time, a mechanism by which the physiological environment can be indicated using *luxABCDE* expression. Focusing on the former, these lux-driven bioluminescent pathogens are commercially available, and used in many localization studies to characterize aspects of infections. The presence of false negatives in our dissemination studies reveals a major concern on current and future literature using these strains, which may be polluted with false negative results. As for the latter, by linking the production of light to the physiology of the bacterial cell, we have shown it is possible to measure the physiological state of the bacterial cell, and, potentially, the carbon source being utilized in particular environments. This phenomenon can be exploited in many ways, but we have demonstrated herein one possible exploitation using antibiotic mechanisms in the application note in Chapter 4. Future innovation will need to focus on determining the practicality of non-*lux*-driven bioluminescence in these catabolite-repressed strains, such as the use of the recently designed toxin-antitoxin (TA) *ffluc*

system [61]. In contrast, in strains containing this lux-driven pathway, future design of metabolic and physiological characterization will provide researchers useful tools for the study of bacterial cells *in vitro* and *in vivo*. For example, physiological activity could be interesting to researchers who study the intracellular survival of pathogens in macrophages, or in other situations where metabolic suppression may be a feature. Quenched bioluminescent light in the presence of other fluorescent dyes for bacterial load, such as the RediJect bacterial detection probe (Lucigen Inc.) could provide molecular imaging research with non-invasive detection of bacterial physiologic response to host immunity or xenobiotics.

In Chapter 4, the Tager 104 genome was probed to determine the identity of virulence factors necessary in the formation of invasive diseases of *S. aureus*, based on molecular imaging data. Based on this genetic background, small nuclear polymorphisms (SNPs) can be determined in these virulence factors and tracked through phylogenetic lineages to provide information on infectivity. This process is similar to that which has been done for vancomycin-intermediate phenotypes [54]. Now that an early-branching, isogenic reference strain which is phenotypically relevant to the invasive infection setting and genotypically 2% similar to the remainder of reference genomes, these whole-genome association studies can take into account the evolution of these virulence factors and the overall contribution to invasive infections.

Taken on the whole, these results demonstrate the use of a molecular imaging/NGS combinatorial technique in the characterization of two major pathogens. The characterization and understanding of the metabolic basis of light production in genetically engineered pathogens, as well as the understanding of the content of a "missing link" in *S. aureus* genomics, provides two major tools for the future advancement towards eradication of these dangerous invasive infections.

References

1. *Centers for Disease Control and Prevention. 2013. Active Bacterial Core Surveillance Report, Emerging Infections Program Network, Group A Streptococcus, 2013.* . 2013.
2. Prevention, C.f.D.C.a., *Antibiotic Resistance Threats in the United States, 2013*. 2013. p. 1-114.
3. Carapetis, J.R., et al., *The global burden of group A streptococcal diseases.* Lancet Infect Dis, 2005. **5**(11): p. 685-94.
4. Cole, J.N., et al., *Molecular insight into invasive group A streptococcal disease.* Nat Rev Microbiol, 2011. **9**(10): p. 724-36.
5. Lamagni, T.L., et al., *Epidemiology of severe Streptococcus pyogenes disease in Europe.* J Clin Microbiol, 2008. **46**(7): p. 2359-67.
6. Abbot, E.L., et al., *Pili mediate specific adhesion of Streptococcus pyogenes to human tonsil and skin.* Cell Microbiol, 2007. **9**(7): p. 1822-33.
7. Cunningham, M.W., *Pathogenesis of group A streptococcal infections.* Clin Microbiol Rev, 2000. **13**(3): p. 470-511.
8. Dale, J.B., et al., *Hyaluronate capsule and surface M protein in resistance to opsonization of group A streptococci.* Infect Immun, 1996. **64**(5): p. 1495-501.
9. Wessels, M.R., et al., *Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci.* Proc Natl Acad Sci U S A, 1991. **88**(19): p. 8317-21.
10. Zinkernagel, A.S., et al., *The IL-8 protease SpyCEP/ScpC of group A Streptococcus promotes resistance to neutrophil killing.* Cell Host Microbe, 2008. **4**(2): p. 170-8.
11. von Pawel-Rammingen, U., Johansson BP, Bjork L., *IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G.* EMBO J., 2002. **21**(7): p. 1607-15.
12. Fernie-King, B.A., D.J. Seilly, and P.J. Lachmann, *The interaction of streptococcal inhibitor of complement (SIC) and its proteolytic fragments with the human beta defensins.* Immunology, 2004. **111**(4): p. 444-52.
13. Fernie-King, B.A., et al., *Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes.* Immunology, 2001. **103**(3): p. 390-8.

14. Zhu, H., et al., *The secreted esterase of group a streptococcus is important for invasive skin infection and dissemination in mice.* Infect Immun, 2009. **77**(12): p. 5225-32.
15. Cole, J.N., et al., *Trigger for group A streptococcal M1T1 invasive disease.* FASEB J, 2006. **20**(10): p. 1745-7.
16. Johannson, H.R., et al., *Osteonecrosis is not a predictor of poor outcomes in primary total hip arthroplasty: a systematic literature review.* Int Orthop, 2011. **35**(4): p. 465-73.
17. McKay, F.C., et al., *Plasminogen binding by group A streptococcal isolates from a region of hyperendemicity for streptococcal skin infection and a high incidence of invasive infection.* Infect Immun, 2004. **72**(1): p. 364-70.
18. Khil, J., et al., *Plasminogen enhances virulence of group A streptococci by streptokinase-dependent and streptokinase-independent mechanisms.* J Infect Dis, 2003. **188**(4): p. 497-505.
19. Cook, S.M., et al., *Streptokinase variants from Streptococcus pyogenes isolates display altered plasminogen activation characteristics - implications for pathogenesis.* Mol Microbiol, 2012. **86**(5): p. 1052-62.
20. Zimlichman, E., et al., *Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system.* JAMA Intern Med, 2013. **173**(22): p. 2039-46.
21. Baddour, L.M., et al., *Infective endocarditis: diagnosis, antimicrobial therapy, and management of complications: a statement for healthcare professionals from the Committee on Rheumatic Fever, Endocarditis, and Kawasaki Disease, Council on Cardiovascular Disease in the Young, and the Councils on Clinical Cardiology, Stroke, and Cardiovascular Surgery and Anesthesia, American Heart Association: endorsed by the Infectious Diseases Society of America.* Circulation, 2005. **111**(23): p. e394-434.
22. Powers, M.E. and J. Bubeck Wardenburg, *Igniting the fire: Staphylococcus aureus virulence factors in the pathogenesis of sepsis.* PLoS Pathog, 2014. **10**(2): p. e1003871.
23. Lowy, F.D., *Staphylococcus aureus infections.* N Engl J Med, 1998. **339**(8): p. 520-32.
24. Yoong, P. and V.J. Torres, *The effects of Staphylococcus aureus leukotoxins on the host: cell lysis and beyond.* Curr Opin Microbiol, 2013. **16**(1): p. 63-9.
25. de Haas, C.J., et al., *Chemotaxis inhibitory protein of Staphylococcus aureus, a bacterial antiinflammatory agent.* J Exp Med, 2004. **199**(5): p. 687-95.

26. Rooijakkers, S.H., et al., *Immune evasion by a staphylococcal complement inhibitor that acts on C3 convertases.* Nat Immunol, 2005. **6**(9): p. 920-7.
27. Rooijakkers, S.H., K.P. van Kessel, and J.A. van Strijp, *Staphylococcal innate immune evasion.* Trends Microbiol, 2005. **13**(12): p. 596-601.
28. Fowler, V.G., Jr., et al., *Clinical identifiers of complicated Staphylococcus aureus bacteremia.* Arch Intern Med, 2003. **163**(17): p. 2066-72.
29. Salgado-Pabon, W., et al., *Superantigens are critical for Staphylococcus aureus Infective endocarditis, sepsis, and acute kidney injury.* MBio, 2013. **4**(4).
30. Panizzi, P., Nahrendorf M, Figueiredo JL, Panizzi JR, Marinelli B, Iwamoto Y, Keliher E, Maddur AA, Waterman P, Kroh HK, Leuschner F, Aikawa E, Swirski FK, Pittet MJ, Hackeng TM, Fuenes-Prior P, Schneedwind O, Bock PE, and Weissleder R. , *In Vitro Detection of Staphylococcus aureus Endocarditis by Targeting Pathogen-Specific Prothrombin Activation.* Nature Medicine, 2011. **17**(9): p. 1142-1146.
31. Friedrich, R., et al., *Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation.* Nature, 2003. **425**(6957): p. 535-9.
32. Panizzi, P., et al., *The staphylocoagulase family of zymogen activator and adhesion proteins.* Cell Mol Life Sci, 2004. **61**(22): p. 2793-8.
33. Panizzi, P., et al., *Fibrinogen substrate recognition by staphylocoagulase.(pro)thrombin complexes.* J Biol Chem, 2006. **281**(2): p. 1179-87.
34. Kroh, H.K., P. Panizzi, and P.E. Bock, *Von Willebrand factor-binding protein is a hysteretic conformational activator of prothrombin.* Proc Natl Acad Sci U S A, 2009. **106**(19): p. 7786-91.
35. Tager, M., *Studies on the coagulase-reacting factor; the reaction of staphylocoagulase with the components of human plasma.* Yale J Biol Med, 1948. **20**(4): p. 369-80.
36. Tager, M., *Concentration, partial purification, properties, and nature of staphylocoagulase.* Yale J Biol Med, 1948. **20**(5): p. 487-501.
37. Tager, M.H., B., *Quantitative Coagulase and Toxin Production by Staphylococcus in Relation to the Clinical Source of the Organisms.* Yale J Biol Med., 1947. **20**(1): p. 41-49.
38. Malachowa, N. and F.R. DeLeo, *Mobile genetic elements of Staphylococcus aureus.* Cell Mol Life Sci, 2010. **67**(18): p. 3057-71.

39. Baba, T., et al., *Genome sequence of Staphylococcus aureus strain Newman and comparative analysis of staphylococcal genomes: polymorphism and evolution of two major pathogenicity islands.* J Bacteriol, 2008. **190**(1): p. 300-10.
40. Harris, S.R., et al., *Evolution of MRSA during hospital transmission and intercontinental spread.* Science, 2010. **327**(5964): p. 469-74.
41. Holt, D.C., et al., *A very early-branching Staphylococcus aureus lineage lacking the carotenoid pigment staphyloxanthin.* Genome Biol Evol, 2011. **3**: p. 881-95.
42. van Oosten, M., et al., *Real-time in vivo imaging of invasive- and biomaterial-associated bacterial infections using fluorescently labelled vancomycin.* Nat Commun, 2013. **4**: p. 2584.
43. Contag, C.H., *Functional Imaging Using Bioluminescent Markers*, in *Molecular Imaging*, R. Weissleder, et al., Editors. 2010, People's Medical Publishing House. p. 118-138.
44. Yhee, J.Y., et al., *Optical imaging of cancer-related proteases using near-infrared fluorescence matrix metalloproteinase-sensitive and cathepsin B-sensitive probes.* Theranostics, 2012. **2**(2): p. 179-89.
45. Francis, K.P., et al., *Visualizing pneumococcal infections in the lungs of live mice using bioluminescent Streptococcus pneumoniae transformed with a novel gram-positive lux transposon.* Infect Immun, 2001. **69**(5): p. 3350-8.
46. Park, H.S., et al., *Membranous cells in nasal-associated lymphoid tissue: a portal of entry for the respiratory mucosal pathogen group A streptococcus.* J Immunol, 2003. **171**(5): p. 2532-7.
47. Niska, J.A., et al., *Monitoring bacterial burden, inflammation and bone damage longitudinally using optical and muCT imaging in an orthopaedic implant infection in mice.* PLoS One, 2012. **7**(10): p. e47397.
48. Guo, Y., et al., *In vivo bioluminescence imaging to evaluate systemic and topical antibiotics against community-acquired methicillin-resistant Staphylococcus aureus-infected skin wounds in mice.* Antimicrob Agents Chemother, 2013. **57**(2): p. 855-63.
49. Treangen, T.J. and S.L. Salzberg, *Repetitive DNA and next-generation sequencing: computational challenges and solutions.* Nat Rev Genet, 2012. **13**(1): p. 36-46.
50. Bashir, A., et al., *A hybrid approach for the automated finishing of bacterial genomes.* Nat Biotechnol, 2012. **30**(7): p. 701-7.

51. Koren, S., et al., *Hybrid error correction and de novo assembly of single-molecule sequencing reads.* Nat Biotechnol, 2012. **30**(7): p. 693-700.
52. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet, 2009. **10**(1): p. 57-63.
53. Howden, B.P., et al., *Reduced vancomycin susceptibility in Staphylococcus aureus, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: resistance mechanisms, laboratory detection, and clinical implications.* Clin Microbiol Rev, 2010. **23**(1): p. 99-139.
54. Alam, M.T., et al., *Dissecting vancomycin-intermediate resistance in staphylococcus aureus using genome-wide association.* Genome Biol Evol, 2014. **6**(5): p. 1174-85.
55. Ralph, A.P. and J.R. Carapetis, *Group a streptococcal diseases and their global burden.* Curr Top Microbiol Immunol, 2013. **368**: p. 1-27.
56. Boxrud, P.D. and P.E. Bock, *Streptokinase binds preferentially to the extended conformation of plasminogen through lysine binding site and catalytic domain interactions.* Biochemistry, 2000. **39**(45): p. 13974-81.
57. Boxrud, P.D. and P.E. Bock, *Coupling of conformational and proteolytic activation in the kinetic mechanism of plasminogen activation by streptokinase.* J Biol Chem, 2004. **279**(35): p. 36642-9.
58. Boxrud, P.D., I.M. Verhamme, and P.E. Bock, *Resolution of conformational activation in the kinetic mechanism of plasminogen activation by streptokinase.* J Biol Chem, 2004. **279**(35): p. 36633-41.
59. Panizzi, P., et al., *Binding of the COOH-terminal lysine residue of streptokinase to plasmin(ogen) kringles enhances formation of the streptokinase.plasmin(ogen) catalytic complexes.* J Biol Chem, 2006. **281**(37): p. 26774-8.
60. Boxrud, P.D., W.P. Fay, and P.E. Bock, *Streptokinase binds to human plasmin with high affinity, perturbs the plasmin active site, and induces expression of a substrate recognition exosite for plasminogen.* J Biol Chem, 2000. **275**(19): p. 14579-89.
61. Loh, J.M. and T. Proft, *Toxin-antitoxin-stabilized reporter plasmids for biophotonic imaging of Group A streptococcus.* Appl Microbiol Biotechnol, 2013. **97**(22): p. 9737-45.
62. Panizzi, P., et al., *In vivo detection of Staphylococcus aureus endocarditis by targeting pathogen-specific prothrombin activation.* Nat Med, 2011. **17**(9): p. 1142-6.

63. Andreu, N., A. Zelmer, and S. Wiles, *Noninvasive biophotonic imaging for studies of infectious disease.* FEMS Microbiol Rev, 2011. **35**(2): p. 360-94.
64. Alam, F.M., et al., *Non-invasive monitoring of Streptococcus pyogenes vaccine efficacy using biophotonic imaging.* PLoS One, 2013. **8**(11): p. e82123.
65. Bergmann, S., M. Rohde, and S. Hammerschmidt, *Glyceraldehyde-3-phosphate dehydrogenase of Streptococcus pneumoniae is a surface-displayed plasminogen-binding protein.* Infect Immun, 2004. **72**(4): p. 2416-9.
66. Seifert, K.N., et al., *Characterization of group B streptococcal glyceraldehyde-3-phosphate dehydrogenase: surface localization, enzymatic activity, and protein-protein interactions.* Can J Microbiol, 2003. **49**(5): p. 350-6.
67. Ringdahl, U., et al., *Molecular co-operation between protein PAM and streptokinase for plasmin acquisition by Streptococcus pyogenes.* J Biol Chem, 1998. **273**(11): p. 6424-30.
68. Wistedt, A.C., et al., *Kringle 2 mediates high affinity binding of plasminogen to an internal sequence in streptococcal surface protein PAM.* J Biol Chem, 1998. **273**(38): p. 24420-4.
69. Wistedt, A.C., et al., *Identification of a plasminogen-binding motif in PAM, a bacterial surface protein.* Mol Microbiol, 1995. **18**(3): p. 569-78.
70. Bergmann, S., H. Schoenen, and S. Hammerschmidt, *The interaction between bacterial enolase and plasminogen promotes adherence of Streptococcus pneumoniae to epithelial and endothelial cells.* Int J Med Microbiol, 2013. **303**(8): p. 452-62.
71. Plow, E.F. and R. Das, *Enolase-1 as a plasminogen receptor.* Blood, 2009. **113**(22): p. 5371-2.
72. Linke, C., et al., *The extracellular protein factor Epf from Streptococcus pyogenes is a cell surface adhesin that binds to cells through an N-terminal domain containing a carbohydrate-binding module.* J Biol Chem, 2012. **287**(45): p. 38178-89.
73. Sanderson-Smith, M.L., et al., *The plasminogen-binding group A streptococcal M protein-related protein Prp binds plasminogen via arginine and histidine residues.* J Bacteriol, 2007. **189**(4): p. 1435-40.
74. Wiles, K.G., et al., *Skizzle is a novel plasminogen- and plasmin-binding protein from Streptococcus agalactiae that targets proteins of human fibrinolysis to promote plasmin generation.* J Biol Chem, 2010. **285**(27): p. 21153-64.

75. McArthur, J.D., et al., *The role of streptokinase as a virulence determinant of Streptococcus pyogenes--potential for therapeutic targeting.* Curr Drug Targets, 2012. **13**(3): p. 297-307.
76. Sazonova, I.Y., et al., *alpha Domain deletion converts streptokinase into a fibrin-dependent plasminogen activator through mechanisms akin to staphylokinase and tissue plasminogen activator.* J Biol Chem, 2004. **279**(24): p. 24994-5001.
77. Svensson, M.D., et al., *Roles of the plasminogen activator streptokinase and the plasminogen-associated M protein in an experimental model for streptococcal impetigo.* Microbiology, 2002. **148**(Pt 12): p. 3933-45.
78. Taylor, S., *Streptokinase; 'streptococcal enzymatic debridement'.* Postgrad Med J, 1950. **26**(297): p. 362-4.
79. Wulf, R.J. and E.T. Mertz, *Studies on plasminogen. 8. Species specificity of streptokinase.* Can J Biochem, 1969. **47**(10): p. 927-31.
80. Sanderson-Smith, M.L., et al., *Bacterial plasminogen receptors: mediators of a multifaceted relationship.* J Biomed Biotechnol, 2012. **2012**: p. 272148.
81. Sanderson-Smith, M.L., et al., *M protein-mediated plasminogen binding is essential for the virulence of an invasive Streptococcus pyogenes isolate.* FASEB J, 2008. **22**(8): p. 2715-22.
82. Sun, H., et al., *Plasminogen is a critical host pathogenicity factor for group A streptococcal infection.* Science, 2004. **305**(5688): p. 1283-6.
83. Agarwal, V., et al., *Streptococcus pneumoniae endopeptidase O (PepO) is a multifunctional plasminogen- and fibronectin-binding protein, facilitating evasion of innate immunity and invasion of host cells.* J Biol Chem, 2013. **288**(10): p. 6849-63.
84. Barman, T.K., et al., *Non invasive real-time monitoring of bacterial infection & therapeutic effect of anti-microbials in five mouse models.* Indian J Med Res, 2011. **134**(5): p. 688-95.
85. Kadioglu, A., et al., *Pneumococcal protein PavA is important for nasopharyngeal carriage and development of sepsis.* Mol Oral Microbiol, 2010. **25**(1): p. 50-60.
86. Kadurugamuwa, J.L., et al., *Reduction of astrogliosis by early treatment of pneumococcal meningitis measured by simultaneous imaging, in vivo, of the pathogen and host response.* Infect Immun, 2005. **73**(12): p. 7836-43.

87. Kadurugamuwa, J.L., et al., *Noninvasive monitoring of pneumococcal meningitis and evaluation of treatment efficacy in an experimental mouse model.* Mol Imaging, 2005. **4**(2): p. 137-42.
88. Kirby, A.C., et al., *SIGNR1-negative red pulp macrophages protect against acute streptococcal sepsis after Leishmania donovani-induced loss of marginal zone macrophages.* Am J Pathol, 2009. **175**(3): p. 1107-15.
89. Malley, R., et al., *Recognition of pneumolysin by Toll-like receptor 4 confers resistance to pneumococcal infection.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1966-71.
90. Mook-Kanamori, B.B., et al., *Daptomycin in experimental murine pneumococcal meningitis.* BMC Infect Dis, 2009. **9**: p. 50.
91. Ogunniyi, A.D., et al., *c-di-GMP is an effective immunomodulator and vaccine adjuvant against pneumococcal infection.* Vaccine, 2008. **26**(36): p. 4676-85.
92. Orihuela, C.J., et al., *Organ-specific models of Streptococcus pneumoniae disease.* Scand J Infect Dis, 2003. **35**(9): p. 647-52.
93. Srivastava, A., et al., *The apoptotic response to pneumolysin is Toll-like receptor 4 dependent and protects against pneumococcal disease.* Infect Immun, 2005. **73**(10): p. 6479-87.
94. Wang, J., et al., *Morphine induces defects in early response of alveolar macrophages to Streptococcus pneumoniae by modulating TLR9-NF-kappa B signaling.* J Immunol, 2008. **180**(5): p. 3594-600.
95. Georgel, P., et al., *A toll-like receptor 2-responsive lipid effector pathway protects mammals against skin infections with gram-positive bacteria.* Infect Immun, 2005. **73**(8): p. 4512-21.
96. Park, H.S. and P.P. Cleary, *Active and passive intranasal immunizations with streptococcal surface protein C5a peptidase prevent infection of murine nasal mucosa-associated lymphoid tissue, a functional homologue of human tonsils.* Infect Immun, 2005. **73**(12): p. 7878-86.
97. Xie, C., et al., *Rescue of impaired fracture healing in COX-2-/- mice via activation of prostaglandin E2 receptor subtype 4.* Am J Pathol, 2009. **175**(2): p. 772-85.
98. Luo, F., et al., *Role of Mga in group A streptococcal infection at the skin epithelium.* Microb Pathog, 2008. **45**(3): p. 217-24.

99. McShan, W.M., et al., *Genome sequence of a nephritogenic and highly transformable M49 strain of Streptococcus pyogenes.* J Bacteriol, 2008. **190**(23): p. 7773-85.
100. Bessen, D.E., et al., *Whole-genome association study on tissue tropism phenotypes in group A Streptococcus.* J Bacteriol, 2011. **193**(23): p. 6651-63.
101. *M9 minimal medium (standard).* 2010: Cold Spring Harbor Protocols.
102. McArthur, J.D., et al., *Allelic variants of streptokinase from Streptococcus pyogenes display functional differences in plasminogen activation.* FASEB J, 2008. **22**(9): p. 3146-53.
103. Zhang, Y., et al., *Characterization of streptokinases from group A Streptococci reveals a strong functional relationship that supports the coinheritance of plasminogen-binding M protein and cluster 2b streptokinase.* J Biol Chem, 2012. **287**(50): p. 42093-103.
104. Kroger, C., et al., *Bistability in myo-inositol utilization by Salmonella enterica serovar Typhimurium.* J Bacteriol, 2011. **193**(6): p. 1427-35.
105. Bessen, D.E. and F.K. Tengra, *Wiring the streptococcal network for alternative lifestyles.* J Infect Dis, 2010. **201**(6): p. 800-2.
106. Li, Z., et al., *Interaction between group A streptococci and the plasmin(ogen) system promotes virulence in a mouse skin infection model.* J Infect Dis, 1999. **179**(4): p. 907-14.
107. Gahan, C.G., *The bacterial lux reporter system: applications in bacterial localisation studies.* Curr Gene Ther, 2012. **12**(1): p. 12-9.
108. Hutchens, M. and G.D. Luker, *Applications of bioluminescence imaging to the study of infectious diseases.* Cell Microbiol, 2007. **9**(10): p. 2315-22.
109. Lee, S.M. and R. Bressler, *Prevention of diabetic nephropathy by diet control in the db/db mouse.* Diabetes, 1981. **30**(2): p. 106-11.
110. Ayala, J.E., et al., *Standard operating procedures for describing and performing metabolic tests of glucose homeostasis in mice.* Dis Model Mech, 2010. **3**(9-10): p. 525-34.
111. Dabul, A.N. and I.L. Camargo, *Clonal complexes of Staphylococcus aureus: all mixed and together.* FEMS Microbiol Lett, 2013.
112. Woods, C.W., et al., *Endocarditis caused by Staphylococcus aureus with reduced susceptibility to vancomycin.* Clin Infect Dis, 2004. **38**(8): p. 1188-91.

113. Meighen, E.A., *Bacterial bioluminescence: organization, regulation, and application of the lux genes.* FASEB J, 1993. **7**(11): p. 1016-22.
114. Meighen, E.A., *Molecular biology of bacterial bioluminescence.* Microbiol Rev, 1991. **55**(1): p. 123-42.
115. Donlan, R.M. and J.W. Costerton, *Biofilms: survival mechanisms of clinically relevant microorganisms.* Clin Microbiol Rev, 2002. **15**(2): p. 167-93.
116. Hall-Stoodley, L., J.W. Costerton, and P. Stoodley, *Bacterial biofilms: from the natural environment to infectious diseases.* Nat Rev Microbiol, 2004. **2**(2): p. 95-108.
117. Kolter, R., D.A. Siegele, and A. Tormo, *The stationary phase of the bacterial life cycle.* Annu Rev Microbiol, 1993. **47**: p. 855-74.
118. Kusters, J.G., et al., *Effects of multiplicity of infection, bacterial protein synthesis, and growth phase on adhesion to and invasion of human cell lines by Salmonella typhimurium.* Infect Immun, 1993. **61**(12): p. 5013-20.
119. Yoshida, K., et al., *The fifth gene of the iol operon of Bacillus subtilis, iolE, encodes 2-keto-myo-inositol dehydratase.* Microbiology, 2004. **150**(Pt 3): p. 571-80.
120. Yoshida, K.I., et al., *Interaction of a repressor and its binding sites for regulation of the Bacillus subtilis iol divergon.* J Mol Biol, 1999. **285**(3): p. 917-29.
121. Davis, R.W.t., et al., *In Vivo Tracking of Streptococcal Infections of Subcutaneous Origin in a Murine Model.* Mol Imaging Biol, 2015.
122. Levering, J., et al., *Role of phosphate in the central metabolism of two lactic acid bacteria--a comparative systems biology approach.* FEBS J, 2012. **279**(7): p. 1274-90.
123. Close, D.M., S. Ripp, and G.S. Sayler, *Reporter proteins in whole-cell optical bioreporter detection systems, biosensor integrations, and biosensing applications.* Sensors (Basel), 2009. **9**(11): p. 9147-74.
124. Hochstadt-Ozer, J. and E.R. Stadtman, *The regulation of purine utilization in bacteria. II. Adenine phosphoribosyltransferase in isolated membrane preparations and its role in transport of adenine across the membrane.* J Biol Chem, 1971. **246**(17): p. 5304-11.
125. Cho, B.K., et al., *The PurR regulon in Escherichia coli K-12 MG1655.* Nucleic Acids Res, 2011. **39**(15): p. 6456-64.

126. Jendresen, C.B., J. Martinussen, and M. Kilstrup, *The PurR regulon in Lactococcus lactis - transcriptional regulation of the purine nucleotide metabolism and translational machinery.* Microbiology, 2012. **158**(Pt 8): p. 2026-38.
127. Fowler, V.G., Jr., et al., *Staphylococcus aureus endocarditis: a consequence of medical progress.* JAMA, 2005. **293**(24): p. 3012-21.
128. Baba, T., et al., *Genome and virulence determinants of high virulence community-acquired MRSA.* Lancet, 2002. **359**(9320): p. 1819-27.
129. Brussow, H., C. Canchaya, and W.D. Hardt, *Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion.* Microbiol Mol Biol Rev, 2004. **68**(3): p. 560-602, table of contents.
130. Plano, L.R., *Staphylococcus aureus exfoliative toxins: how they cause disease.* J Invest Dermatol, 2004. **122**(5): p. 1070-7.
131. Sperber, W.H. and S.R. Tatini, *Interpretation of the tube coagulase test for identification of Staphylococcus aureus.* Appl Microbiol, 1975. **29**(4): p. 502-5.
132. Tager, M. and M.C. Drummond, *Staphylocoagulase.* Ann N Y Acad Sci, 1965. **128**(1): p. 92-111.
133. Read, T.D. and S.W. Satola, *Using genomics to standardize population analysis profile-area under the curve ratio for vancomycin-intermediate Staphylococcus aureus.* J Clin Microbiol, 2014. **52**(10): p. 3824-6.
134. Tager, M. and A.L. Lodge, *Influence of the physiological blood clotting process on the coagulation of blood by staphylocoagulase.* J Exp Med, 1951. **94**(1): p. 73-85.
135. Tager, M. and A.L. Lodge, *Changes in the properties of the coagulase-reacting factor of plasma after separation from prothrombin by Seitz filtration.* J Immunol, 1951. **67**(1): p. 63-9.
136. Panizzi, P., et al., *Novel fluorescent prothrombin analogs as probes of staphylocoagulase-prothrombin interactions.* J Biol Chem, 2006. **281**(2): p. 1169-78.
137. Panizzi, P., J.R. Stone, and M. Nahrendorf, *Molecular Imaging of Endocarditis.* Journal of Nuclear Cardiology, 2014.
138. Davis, R., et al., *Complete Genome Sequence of Staphylococcus aureus Tager 104, a Sequence Type 49 Ancestor.* Genome Announc, 2013. **1**(5).

139. Chin, C.S., et al., *Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data.* Nat Methods, 2013. **10**(6): p. 563-9.
140. Hunt, M., et al., *A comprehensive evaluation of assembly scaffolding tools.* Genome Biol, 2014. **15**(3): p. R42.
141. Aziz, R.K., et al., *The RAST Server: rapid annotations using subsystems technology.* BMC Genomics, 2008. **9**: p. 75.
142. Brettin, T., et al., *RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes.* Sci Rep, 2015. **5**: p. 8365.
143. Overbeek, R., et al., *The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST).* Nucleic Acids Res, 2014. **42**(Database issue): p. D206-14.
144. Nurk, S., et al., *Assembling single-cell genomes and mini-metagenomes from chimeric MDA products.* J Comput Biol, 2013. **20**(10): p. 714-37.
145. Kurtz, S., et al., *Versatile and open software for comparing large genomes.* Genome Biol, 2004. **5**(2): p. R12.
146. Enright, M.C., et al., *Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus.* J Clin Microbiol, 2000. **38**(3): p. 1008-15.
147. Vesth, T., et al., *CMG-biotools, a free workbench for basic comparative microbial genomics.* PLoS One, 2013. **8**(4): p. e60120.
148. Darling, A.E., B. Mau, and N.T. Perna, *progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement.* PLoS One, 2010. **5**(6): p. e11147.
149. Bhat, M., et al., *Staphylococcus aureus ST398, New York City and Dominican Republic.* Emerg Infect Dis, 2009. **15**(2): p. 285-7.
150. Uhlemann, A.C., et al., *Identification of a highly transmissible animal-independent Staphylococcus aureus ST398 clone with distinct genomic and cell adhesion properties.* MBio, 2012. **3**(2).
151. Nubel, U., et al., *A timescale for evolution, population expansion, and spatial spread of an emerging clone of methicillin-resistant Staphylococcus aureus.* PLoS Pathog, 2010. **6**(4): p. e1000855.

152. Stegger, M., et al., *Genome sequence of Staphylococcus aureus strain 11819-97, an ST80-IV European community-acquired methicillin-resistant isolate.* J Bacteriol, 2012. **194**(6): p. 1625-6.
153. Gill, S.R., et al., *Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain.* J Bacteriol, 2005. **187**(7): p. 2426-38.
154. Lindqvist, M., et al., *Detection and characterisation of SCCmec remnants in multiresistant methicillin-susceptible Staphylococcus aureus causing a clonal outbreak in a Swedish county.* Eur J Clin Microbiol Infect Dis, 2012. **31**(2): p. 141-7.
155. Ben Zakour, N.L., et al., *Genome-wide analysis of ruminant Staphylococcus aureus reveals diversification of the core genome.* J Bacteriol, 2008. **190**(19): p. 6302-17.
156. Guinane, C.M., et al., *Evolutionary genomics of Staphylococcus aureus reveals insights into the origin and molecular basis of ruminant host adaptation.* Genome Biol Evol, 2010. **2**: p. 454-66.
157. Holden, M.T., et al., *A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic.* Genome Res, 2013. **23**(4): p. 653-64.
158. Sieradzki, K., et al., *Evolution of a vancomycin-intermediate Staphylococcus aureus strain in vivo: multiple changes in the antibiotic resistance phenotypes of a single lineage of methicillin-resistant S. aureus under the impact of antibiotics administered for chemotherapy.* J Clin Microbiol, 2003. **41**(4): p. 1687-93.
159. Howden, B.P., et al., *Complete genome sequence of Staphylococcus aureus strain JKD6008, an ST239 clone of methicillin-resistant Staphylococcus aureus with intermediate-level vancomycin resistance.* J Bacteriol, 2010. **192**(21): p. 5848-9.
160. Chua, K., et al., *Complete genome sequence of Staphylococcus aureus strain JKD6159, a unique Australian clone of ST93-IV community methicillin-resistant Staphylococcus aureus.* J Bacteriol, 2010. **192**(20): p. 5556-7.
161. Garcia-Alvarez, L., et al., *Meticillin-resistant Staphylococcus aureus with a novel mecA homologue in human and bovine populations in the UK and Denmark: a descriptive study.* Lancet Infect Dis, 2011. **11**(8): p. 595-603.

162. Huang, T.W., et al., *Complete genome sequence of Staphylococcus aureus M013, a pvl-positive, ST59-SCCmec type V strain isolated in Taiwan.* J Bacteriol, 2012. **194**(5): p. 1256-7.
163. Holden, M.T., et al., *Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9786-91.
164. Hiramatsu, K., et al., *Dissemination in Japanese hospitals of strains of Staphylococcus aureus heterogeneously resistant to vancomycin.* Lancet, 1997. **350**(9092): p. 1670-3.
165. Liu, C. and H.F. Chambers, *Staphylococcus aureus with heterogeneous resistance to vancomycin: epidemiology, clinical significance, and critical assessment of diagnostic methods.* Antimicrob Agents Chemother, 2003. **47**(10): p. 3040-5.
166. Kuroda, M., et al., *Whole genome sequencing of meticillin-resistant Staphylococcus aureus.* Lancet, 2001. **357**(9264): p. 1225-40.
167. Herbert, S., et al., *Repair of global regulators in Staphylococcus aureus 8325 and comparative analysis with other clinical isolates.* Infect Immun, 2010. **78**(6): p. 2877-89.
168. Herron-Olson, L., et al., *Molecular correlates of host specialization in Staphylococcus aureus.* PLoS One, 2007. **2**(10): p. e1120.
169. Fitzgerald, J.R., et al., *Fine-structure molecular epidemiological analysis of Staphylococcus aureus recovered from cows.* Epidemiol Infect, 1997. **119**(2): p. 261-9.
170. Schijffelen, M.J., et al., *Whole genome analysis of a livestock-associated methicillin-resistant Staphylococcus aureus ST398 isolate from a case of human endocarditis.* BMC Genomics, 2010. **11**: p. 376.
171. Li, Y., et al., *Complete genome sequence of Staphylococcus aureus T0131, an ST239-MRSA-SCCmec type III clone isolated in China.* J Bacteriol, 2011. **193**(13): p. 3411-2.
172. Diep, B.A., et al., *Emergence of multidrug-resistant, community-associated, methicillin-resistant Staphylococcus aureus clone USA300 in men who have sex with men.* Ann Intern Med, 2008. **148**(4): p. 249-57.
173. Diep, B.A., et al., *Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus.* Lancet, 2006. **367**(9512): p. 731-9.

174. Highlander, S.K., et al., *Subtle genetic changes enhance virulence of methicillin resistant and sensitive Staphylococcus aureus.* BMC Microbiol, 2007. **7**: p. 99.
175. Gonzalez, B.E., et al., *Severe Staphylococcal sepsis in adolescents in the era of community-acquired methicillin-resistant Staphylococcus aureus.* Pediatrics, 2005. **115**(3): p. 642-8.
176. Schaaff, F., A. Reipert, and G. Bierbaum, *An elevated mutation frequency favors development of vancomycin resistance in Staphylococcus aureus.* Antimicrob Agents Chemother, 2002. **46**(11): p. 3540-8.
177. Feil, E.J., et al., *eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data.* J Bacteriol, 2004. **186**(5): p. 1518-30.
178. Mwangi, M.M., et al., *Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing.* Proc Natl Acad Sci U S A, 2007. **104**(22): p. 9451-6.
179. Akhtar, M.S., et al., *Antimicrobial peptides as infection imaging agents: better than radiolabeled antibiotics.* Int J Pept, 2012. **2012**: p. 965238.
180. Kos, V.N., et al., *Comparative genomics of vancomycin-resistant Staphylococcus aureus strains and their positions within the clade most commonly associated with Methicillin-resistant S. aureus hospital-acquired infection in the United States.* MBio, 2012. **3**(3).
181. Overesch, G., et al., *The increase of methicillin-resistant Staphylococcus aureus (MRSA) and the presence of an unusual sequence type ST49 in slaughter pigs in Switzerland.* BMC Vet Res, 2011. **7**: p. 30.
182. Paterson, G.K., et al., *The newly described mecA homologue, mecALGA251, is present in methicillin-resistant Staphylococcus aureus isolates from a diverse range of host species.* J Antimicrob Chemother, 2012. **67**(12): p. 2809-13.
183. Winstel, V., et al., *Wall teichoic acid structure governs horizontal gene transfer between major bacterial pathogens.* Nat Commun, 2013. **4**: p. 2345.