New Genomics Tools and Strategies for Studying Antibiotics and Antibiotic-Resistance in Staphylococcus Aureus

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Santiago, Marina Joy. 2016. New Genomics Tools and Strategies for Studying Antibiotics and Antibiotic-Resistance in Staphylococcus Aureus. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:33493460
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

New Genomics Tools and Strategies for Studying Antibiotics and Antibiotic-resistance in

Staphylococcus aureus

A dissertation presented

by

Marina Joy Santiago

to

The Committee on Higher Degrees in Chemical Biology

In partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge, Massachusetts

April 27, 2016

New genomics tools and strategies for studying antibiotics and antibiotic-resistance in
Staphylococcus aureus

Abstract

Staphylococcus aureus is a gram positive coccoid pathogen that causes intractable infections in hospitals and communities around the world, and tens of thousands of people die of these infections every year. In order to combat these antibiotic-resistant infections, we need to better understand the genes involved in resistance to the cell stress caused by antibiotic treatment, which will enable the discovery of new antimicrobials and the development of novel therapeutic strategies. We chose to use an approach to this problem that utilizes a new phage-based high frequency of transposition system. In this work, we adapted this system so that transposon mutant libraries can be made and sequenced using next-generation sequencing (NGS) in any strain of *S. aureus.* We validated our new platform by performing a temperature screen and identifying mutants that are significantly resistant or sensitive to temperature-stress. Next, we created transposon libraries in two MRSA strains to show that this system can be broadly applied to other *S. aureus* strains, and we used one of these libraries to identify a new interaction between two genes involved in the secretion of sortase-anchored surface proteins. To better understand antibiotic-resistance, we performed Tn-Seq on transposon libraries treated with a small panel of six different antibiotics to identify intrinsic resistance factors to these antibiotics. We identified two new intrinsic resistance factors, *SAOUHSC_01025* and *SAOUHSC_01050,* that sensitize to many cell envelope targeting antibiotics and may be involved in hemolysin regulation. Finally, we expanded this approach to sequence transposon libraries treated with 25 different antibiotics. Based on these data, we were able to develop methods for predicting the mechanism of action of new antibiotics. These methods involve

identifying genes upregulated by transposon insertion and applying machine learning algorithms to identify similarities to a curated panel of well-studied antibiotics with known mechanisms of action. This work will enable many new functional genomics studies in *S. aureus*, and it will allow us to gain a better understanding of antibiotic resistance in this dangerous pathogen.

Table of Contents

Acknowledgements

I would like to thank:

Suzanne Walker for giving me the opportunity to learn so many new things with this project;

All the members of the Walker lab for their support,

Especially Mithila Rajagopal for being a great partner on this project and an amazing friend;

Julia Rogers and Brian Fults for being the best friends in the world;

My family;

and

Nestle Toll House Semi-Sweet Chocolate Chips

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Antibiotic resistant infections

Antibiotics and Antibiotic-resistance

Antibiotics have been in the clinic so long they are often taken for granted. The first widely-used antimicrobials were the sulfonamides, which were introduced in the 1930's (1), but it was the discovery of the first β-lactam antibiotic, penicillin, by Alexander Fleming in 1928 (2), and its mass production in the 1940's that began the golden era of antibiotic discovery (3). Many new antibiotics were discovered by pharmaceutical companies. Most of these new antibiotics were natural products that were extracted from bacteria or fungi in much the same way that penicillin was (3). For many years after the discovery of penicillin, the vast majority of bacterial infections have been treatable with these antibiotics.

More recently, the situation has changed. Resistance to antibiotics has become very common, and now resistance in the clinic has been observed to every class of antibiotics (Figure 1) (4). Antibiotic resistance can develop through the evolution of new functions or through the acquisition of resistance factors through horizontal gene transfer (5). There are three major classes of antibiotic resistance (6): 1) Resistance through the increased export of the antibiotic by multi-drug resistance pumps, which prevent the antibiotic from reaching its target (ex: NorA is a pump whose expression is upregulated with ciprofloxacin resistance (7)); 2) resistance through the modification of the structure of the antibiotic (by adding moieties or by breaking the antibiotic apart), which prevents the antibiotic from properly binding its target (ex: FosB can hydrolyze the epoxide ring of fosfomycin, rendering it inactive (8)); 3) Resistance through some structural change in the antibiotic's target which prevents the antibiotic from properly binding it (ex: PBP2a is a transpeptidase with a low affinity for β-lactam antibiotics, which prevents it from becoming inactivated (9)).

1
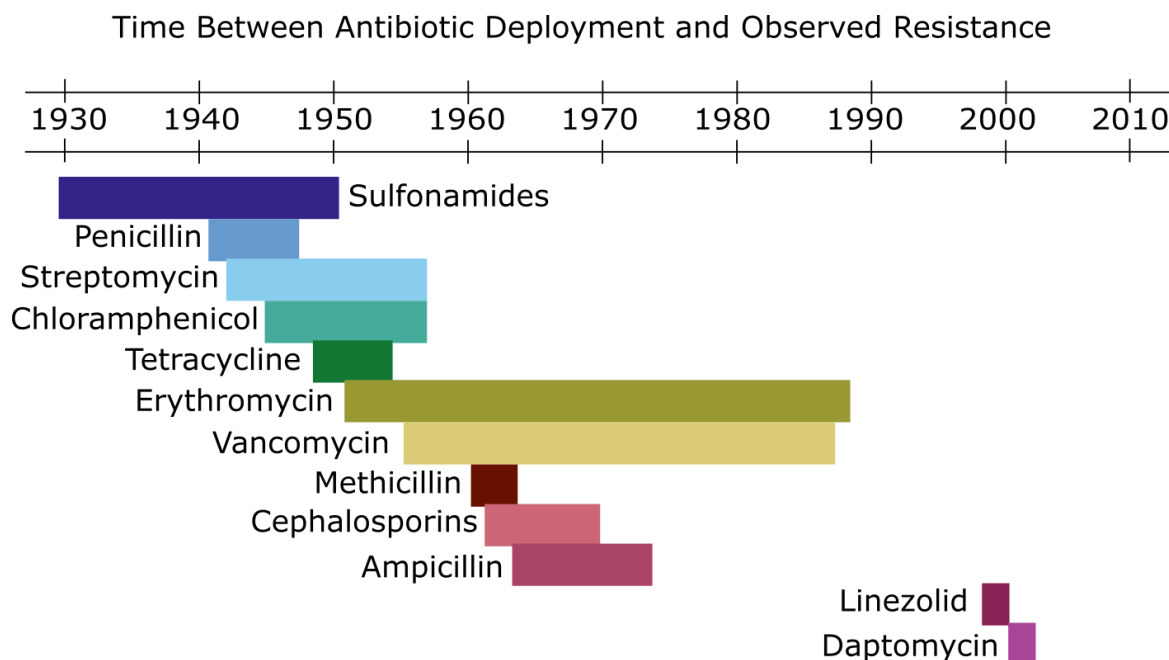
Figure 1. Antibiotic-resistance has been observed to every class of antibiotic (Adapted from Clatworthy et al. 2007 Nature Chemical Biology – Fig1 (4)). These antibiotics belong to some of the major clinical antibiotic classes. These classes include the sulfonamides which inhibit folate synthesis by binding dihydropteroate synthase (10), the β-lactams (penicillin, methicillin, cephalosporins, and ampicillin) which inhibit peptidoglycan biosynthesis by binding the peptidoglycan biosynthetic proteins (PBPs) (11), the aminoglycosides (streptomycin) which perturbs the 30S unit of the ribosome resulting in inaccurate mRNA translation (12, 13), the macrolides (erythromycin) which also inhibit protein synthesis by binding to the 50S ribosomal subunit (14), the glycopeptides (vancomycin) which inhibit peptidoglycan synthesis by binding to and sequestering peptidoglycan precursors (15, 16), the oxazolidinones (linezolid) which binds to the 50S subunit of the ribosome preventing the protein translation initiation complex (17), the lipopeptides (daptomycin) which disrupt the membrane and delocalize peptidoglycan biosynthetic machinery (18, 19), and the tetracyclines bind to the 30S ribosomal subunit (20) while chloramphenicol binds to the 50S ribosomal subunit (21). **Continued page 3.**

Figure 1 continued. Each bar begins when an antibiotic was deployed in the clinic. The bar ends when clinical resistance was observed. This figure also highlights the Golden Age of antibiotic development in the 1940's through the 1960's as well as the dearth of new antibiotics in the 1970's, 1980's, and 1990's(22).

One of the reasons why antibiotic-resistance has become a major problem recently is that, from the 1970s through the 1990s, the antibiotic discovery pipeline almost completely dried up (23). This is not only because the low-hanging fruit (abundant natural products from easily cultured microbes) had been picked, which made the discovery of new antibiotics more challenging, but also because the economics and regulatory environment of antibacterial development made it difficult for newly-discovered antibiotics to come to market (24). The Food and Drug Administration in the United States had increased the requirements for bringing an antibiotic through the clinical trial process, which increased the cost of developing an antibiotic. Furthermore, antibiotics for humans do not make much money for pharmaceutical companies. This is because, unlike a cardiovascular drug the patient does not take antibiotics regularly for years of their life, and unlike cancer drugs, antibiotics are relatively inexpensively priced. Also, some antibiotics are narrow spectrum, and are only useful for certain classes or even species of bacteria. That means that a new antibiotic might not have a large number of people for whom it will be useful, and if it is a good antibiotic, the patient will not be taking it for a very long time (24). Moreover, high-quality rapid diagnostics are needed along with the antibiotic if a narrow-spectrum antibiotic is to be very effective at treating patients (25). Having to develop a diagnostic to go along with an antibiotic further increases the cost of bringing a drug to market (25). If pharmaceutical companies increase the cost of an antibiotic in order to make up for these losses, patients and doctors rebel against what appears to be an unnecessarily-expensive drug. In order to make money on antibiotics, pharmaceutical companies may market their

antibiotic as a growth promoter in for livestock, which increases the likelihood that resistance to that antibiotic will develop (24). A vast shift in how we develop and use new antibiotics is needed to encourage major pharmaceutical companies to re-enter this field (26).

***Staphylococcus aureus* is a dangerous pathogen**

In 2008, the Infectious Diseases Society of America highlighted a group of pathogens that have the ability to "escape" normal antibiotic treatment. These bacteria, called the ESKAPE pathogens, are part of a new paradigm of antibiotic resistance that will require novel strategies to treat. The ESKAPE pathogens consist of *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumanii, Pseudomonas aeruginosa,* and *Enterobacter* species (27). All of these are organisms of concern, but we are particularly interested in *S. aureus* infections because these infections can be highly-resistant to antibiotics and incredibly deadly. Approximately 2 million people in the United States fall ill with antibiotic-resistant infections every year, resulting in 23,000 deaths. Though invasive MRSA infections only account for 80,000 of the 2 million infections, this organism causes almost half, or 11,000, of the 23,000 deaths (28).

Staphylococcus aureus is a gram-positive coccoid species of bacteria. It is a common human commensal, colonizing the skin, especially the nose and nares (29). Approximately 20-30% of Americans are colonized by *S. aureus* without any ill effects (29, 30). However, some strains of *S. aureus* are pathogenic and can cause cutaneous abscesses that, if untreated, can lead to pneumonia, sepsis, necrotizing fasciitis, and death (31). Some *S. aureus* strains can be very virulent without having any antibiotic-resistance factors. *S. aureus* Newman, a classic example of one of these strains, is a virulent, but antibiotic-sensitive, strain isolated from a human infection in 1952 (32). At least thirty genes in this strain were shown to be important for *S. aureus* pathogenesis, and these genes are shared among many different strains (33). The virulence of *S. aureus* has been shown to be due to secreted enzymes and toxins such as

4

coagulases, hemolysins, toxic shock syndrome toxin (TSST-1), and Panton-Valentine leucocidin (34-36). *S. aureus* also produces staphyloxanthin, an antioxidant that protects the bacteria from reactive oxygen species (37) and allows the cell to survive longer in the presence of neutrophils (38).

Both methicillin-sensitive *S. aureus* (MSSA), and methicillin-resistanct *S. aureus* (MRSA) strains cause dangerous infections, but MRSA strains are also resistant to β-lactam antibiotics. MRSA strains can survive very high concentrations of β-lactam antibiotics because they contain the SCC*mec* cassette, a mobile genetic element encoding a penicillin-binding protein (PBP2a) with a low affinity for β-lactam antibiotics as well as the genes that regulate its expression (Figure 2) (9). β-lactam antibiotics function by binding to and inactivating the transpeptidase domain of penicillin-binding proteins (PBPs), which crosslink peptidoglycan strands to synthesize the cell wall (11). The expression of PBP2a permits the cell to continue to synthesize peptidoglycan, allowing it to grow and divide even when the other PBPs are inactivated by a β-lactam antibiotic (9).
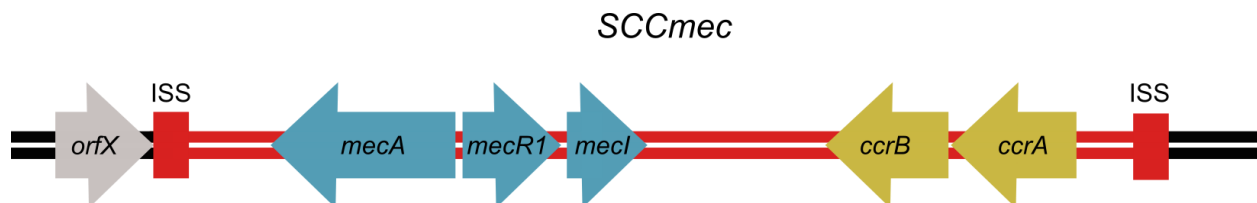
SCCmec

Figure 2. The SCC*mec* mobile genetic element encodes β-lactam resistance. There are eight classes of the SCC*mec* element (shown in red), but they all encode some version of these genes (39). The SCC*mec* element is surrounded by Insertion Site Sequences (ISS) depicted as red rectangles which control where it inserts into the genome, at the 3' end of *orfX*. *ccrA* and *ccrB* are responsible or integration and excision of the SCC*mec* element. *mecA* encodes PBP2a, the β-lactam resistant PBP, while *mecR1* and *mecI* regulate the expression of PBP2a (39).

There are many types of MRSA strains. MRSA strains are broadly categorized as healthcare-associated (HA-MRSA) or community-acquired (CA-MRSA). HA-MRSA infections are one of the most common healthcare-associated infections, but better monitoring of hospitals and other clinical interventions have successfully decreased the number of HA-MRSA infections over the last ten years (40, 41). However, the number of CA-MRSA infections has remained constant (42, 43). In fact, one of the most common strains of MRSA in the United States, USA300, is a CA-MRSA strain (44). Other worrisome strains of MRSA include vancomycin intermediate resistant *S. aureus* (VISA) and vancomycin resistant *S. aureus* (VRSA) (28). Vancomycin, which kills cells by binding to a dipeptide found in peptidoglycan precursors, preventing new peptidoglycan synthesis, is considered to be an antibiotic of last resort for *S. aureus* infections (45). VRSA strains have acquired the vancomycin resistance gene *vanA,* likely via horizontal gene transfer from *Enterococcus faecalis.* This gene confers high level resistance to vancomyin by changing the terminal amino acid residues on the stem peptide of peptidoglycan precursors D-ala-D-ala to D-ala-D-lactate (46). VRSA infections are currently not very common (between 2002 and 2013, only 13 cases were conclusively identified), but VISA infections are much more common (28, 47). Since 1997, they have been identified in Japan, the United Kingdom, France, the United States, and Brazil. These strains are less resistant to vancomycin than VRSA, but are more resistant than *S. aureus* and MRSA strains (48). It is impossible to treat VISA infections with vancomycin because at higher concentrations, vancomycin toxicity to the patient becomes a problem (49). Many types of mutations can confer the VISA phenotype, but a large number are in cell stress sensing regulatory systems which cause thickening of the cell wall, preventing vancomycin from reaching lipid II at the cell septum where new peptidoglycan biosynthesis is occurring (50).

Antibiotic resistance in *S. aureus*

   S. aureus does not have high-level resistance to all antibiotics, but even MSSA strains are well known to be highly adaptable to stressful conditions like antibiotic treatment. There are mechanisms for up- or down-regulating the expression of genes to promote the cell's survival. Different genes are required for resistance to different antibiotics. Genes that become essential in the presence of antibiotic but are not acquired resistance genes are termed intrinsic resistance factors. For example, *mprF* is a lysylphosphatidylglycerol synthase (51) that is normally non-essential, but becomes essential during treatment with daptomycin (52). Daptomycin acts by binding to at least one calcium ion and inserting into the membrane in a phosphatidyl-glycerol dependent manner, depolarizing it and delocalizing the cell division machinery (18, 19, 53). Upregulation or activation of *mprF* results in an increased resistance to daptomycin (54, 55). Mechanisms like this exist for many other gene and antibiotic pairs. A better understanding of the intrinsic resistance factors could expose some of *S. aureus*'s vulnerabilities, which could be targeted and lead to new strategies for antibiotic development.

   Another form of antibiotic resistance in MRSA is the small colony variant (SCV) phenotype. Infections by SCV strains are persistent and often recurrent (56, 57). SCV infections are difficult to treat because they are highly antibiotic-resistant and difficult to isolate from an infection (58). These strains come about as a result of defects in the electron transport chain, often due to a mutation in the menaquinone or cytochrome biosynthetic pathways that causes a switch from oxidative to fermentative growth (59, 60). Slow growth can cause a decrease in susceptibility to antibiotics because the cells are not very metabolically active. Furthermore, these mutations can cause a decrease in membrane potential, which is required for the uptake of aminoglycoside antibiotics and cationic antimicrobial peptides (59). Therefore, these antibiotics are ineffective against SCV infections.

MRSA can cause some of the most dangerous infections because it is capable of responding rapidly and effectively to the stress caused by antibiotic treatment. And though everyday more bacteria evolve resistance to antibiotics, very few new antibiotics are being developed for treating these infections. These new antibiotics are generally chemical modifications of antibiotics currently in use. Ceftaroline and ceftobiprole are fifth generation cephalosporins, dalbavancin is a teicoplanin derivative, and tedizolid is another oxazolidinone. Or, they are combinations of currently-approved drugs combined with a potentiating compound. Examples of these include ceftazidime-avibactam and ceftolozane-tazobactam (61). Since the FDA approval of daptomycin in 2003, an antibiotic with a new chemical scaffold has not been developed (62). If we are to effectively combat antibiotic resistant infections, we must develop new tools for studying *S. aureus* as well as creative strategies for treating these infections. The tool we chose to use for studying *S. aureus* is the transposon mutant library. A history of how these tools have been used to better understand bacterial physiology follows.

1.2. History of transposons as genetic tools

What is a transposon?

Transposons were discovered by Barbara McClintock in 1928 (63), but it was not until 1950 after DNA was better understood that the scientific community comprehended their importance (64). Before their discovery, DNA was thought to be an unchangeable and linear set of genes on a chromosome. However, while studying chromosome breakage and fusion in maize, she identified two loci that that could change position (64). These were the first transposons, a class of mobile genetic elements. With the help of a transposase or other cellular factors, they can jump to different areas of the genome (65, 66). Transposons are incredibly common. They have been identified in all kingdoms of life, and they make up a large fraction of the 66% of the human genome that is made up of repetitive sequences (67, 68). In bacteria, transposons and other mobile genetic elements are associated with the acquisition of genes that increase their fitness such as virulence factors and antibiotic-resistance cassettes (69).

There are many different kinds of transposons. Class I transposons, or the retrotransposons, use an RNA-intermediate to move to a new location in the genome. Class II transposons encode a transposase that recognizes the transposon's inverted terminal repeats, allowing the enzyme to "cut and paste" the transposon into a new location in the genome (70). In addition, transposons can integrate into specific sites in the genome or they may integrate randomly (71). Site-specific transposons require homology between the original and new site of insertion. These homologous sequences can vary from the AT-rich pentanucleotide sequences preferred by *Helicobacter pylori*'s IS*605* element (72) to the 17-bp identical sequences used by the SXT transposon of *Virbrio cholera* (73). Other transposons can insert randomly throughout the genome. This class includes the phage-derived transposons which such as the Mu transposon (74) and the *Drosophila*-derived transposon, *mariner* (75). The *mariner* transposon

has been widely used for many genetics and genomics studies because of the ease in adapting it for various organisms and platforms (76). It can insert randomly into any TA dinucleotide site in the genome. This type of transposon may also someday be used for gene therapy in humans and was awarded Molecule of the Year in 2009 by The International Society for Molecular and Cell Biology and Biotechnology Protocols and Researches (77).

Transposon applications in bacteria

Randomly-integrated transposons have been used for a variety of traditional applications including genetic footprinting, creating transcriptional and translational fusions, DNA sequencing, scanning-linker mutagenesis, signature-tagged mutagenesis, and transposon site hybridization. Genetic footprinting predicts whether a gene is essential in a condition of interest by performing PCR using a pair of primers that anneal to the transposon and the gene of interest after growth in a specific condition (78). Transcriptional and translational fusion libraries can be created by placing promoterless reporters (ex: *lacZ, luxAB, xylE*) inside a transposon and then creating the transposon mutant library library (79). This allows one to screen for genes that are differentially regulated in response to an environmental stimulus (80). DNA and cDNA sequencing can be performed using a transposon mutant library by using transposon-specific primers for sequencing the genomic regions next to the transposon insertion and assembling those segments into the full sequence (81). Scanning linker mutagenesis is a powerful tool for structure-function studies. This technique involves using transposons in create in-frame insertion mutants in a gene. Then, site-specific recombination removes the unnecessary DNA segments of the transposon itself, leaving only a small insertion in the gene (76, 82). Signature-tagged mutagenesis involves making a transposon mutant library from a transposon containing different random sequences used as signature tags. This was an important development because it allowed for the simultaneous screening for multiple genes important for growth in a specific condition (83). It has traditionally been used for assessment of genes required for

10

virulence during infection (84, 85). Transposon-site hybridization (TraSH) uses a custom microarray to identify conditionally-essential genes in a high-throughput manner (86). In short, fluorescently-labeled probes derived from the transposon-genome junctions are created from a transposon mutant library and hybridized to an array containing every gene in the genome. Changes in fluorescence at each spot on the array after growth in a condition of interest are identified, and these correspond to transposon mutants with a change in fitness in that condition (86-88).

Transposon mutagenesis and next-generation sequencing (NGS)

As the cost of DNA sequencing has decreased, high-throughput methods similar to signature-tagged mutagenesis and TraSH have become more common. Numerous tools and techniques for next-generation sequencing (NGS) of transposon mutant libraries have been created and published, but each follows a similar series of steps (89). For each, a transposon library is created and grown in a condition of interest. Because of the danger of antibiotic-resistant bacterial infections, these conditions are often antibiotic treatment or growth in a model organism (90-95), but others have investigated the impact of other kinds of stress such as uranium stress (96) and changes in CO_2 concentration (97). After growth in the condition of interest, the cells are collected, genomic DNA is isolated, and transposon-genome junctions are purified away from the rest of the DNA. The location of the transposon insertion is sequenced using NGS, and genes important for fitness in that condition are found by identifying genes with significant changes in numbers of reads that map to them (Figure 3).
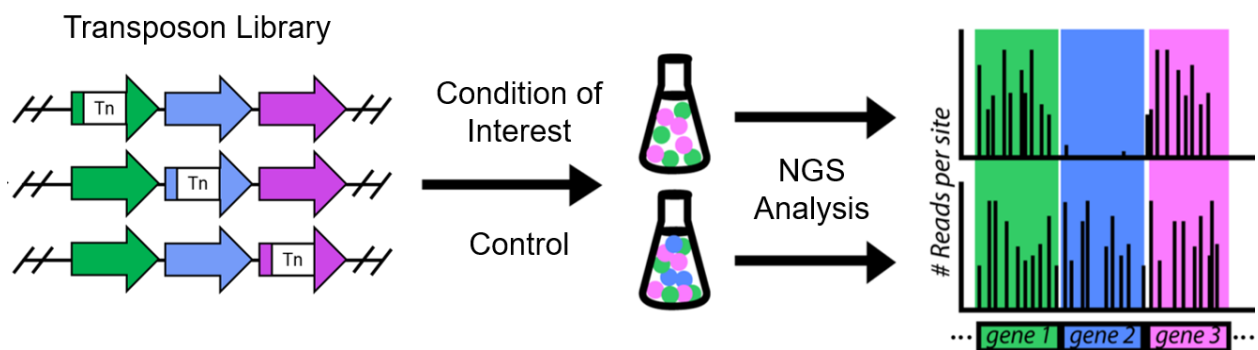
Figure 3. Transposon libraries can be sequenced with NGS to discover genes required for survival in a condition of interest. To perform these experiments, a transposon mutant library is grown in the condition of interest (ex: antibiotic treatment) and a control condition. Cells with transposon insertions into genes which are required for growth in the condition of interest will die and drop out of the library (ex: blue cells). Once the cells have been collected, the DNA extracted, and transposon insertion locations identified using NGS, we can identify genes with significantly fewer reads mapping to them (ex: blue gene). This allows us to identify genes required for growth in a given condition.

Tn-Seq, TraDIS, INSeq, and HITS are four similar techniques for sequencing the location of transposon insertions(89). They differ only in the method used to prepare transposon-genome junctions for DNA sequencing. Tn-Seq and InSeq take advantage of the MmeI restriction enzyme which cuts 20bp down from its recognition site (98, 99). With one base pair change, the MmeI recognition site can be incorporated near the end of the transposon's inverted terminal repeats. This allows capture of the transposon as well as ~16bp of genomic DNA next to it, allowing mapping of the transposon insertion site to the *S. aureus* reference genome. A major difference between the Tn-Seq and INSeq protocol is a step in the INSeq protocol that uses a linear PCR with a biotinylated primer and purification with streptavidin beads. This step decreases the amount of DNA, enzymes, and other reagents required to

prepare the DNA for NGS. In addition, it may increase the sensitivity of the resulting sequencing data (89, 99). HITS and TraDIS do not use MmeI digestion to purify the transposon-genome junctions, but instead use DNA shearing (100, 101). After shearing, the ends of the DNA are repaired and a poly(A) overhang is added. In both, PCR is used to enrich for Tn-chromosome junctions, but while the PCR products are directly sequenced in the TraDIS method (101), HITS uses a biotinylated primer for the PCR, size selection by running the PCR product on a gel, and affinity purification of the DNA, before sequencing with NGS (100). All of these techniques have been widely used and broadly-applied to many different bacterial species to study answer different biological questions (89), though these techniques are all now largely referred to as "Tn-Seq" by the field.

As these techniques have become more popular, there has been an effort to make the analysis of NGS data more accessible to the average biologist who has little computational experience. High-quality software is important for making these techniques widely available, and initially, most labs wrote their own custom scripts to understand the data. The best way of normalizing experimental data and assessing statistical significance likely depends on the characteristics of the transposon library, the experiment performed, and the question the transposon sequencing data is hoping to answer. Some of the custom methods for analysis of transposon sequencing data include using a Bayesian analysis (102) or a Hidden-Markov Model (103, 104) for predicting gene essentiality. Our lab has previously used the Mann-Whitney statistical test for identifying conditionally essential genes after growth in a non-lethal compound (95, 105). Furthermore, the Tufts Galaxy Server has its own custom Tn-Seq analysis pipeline which can be used by anyone sequencing transposon libraries through their sequencing facility (http://galaxy.med.tufts.edu/).

One of the first "easy-to-use" programs published was ESSENTIALS, which did not require any knowledge of computing languages to run. This online software identifies essential

and conditionally-essential genes and can be used with single or multiple transposon libraries (106). However, other pipelines and software packages have since been published. These include the ARTIST package which combines the Hidden-Markov Model for predicting essential genes (EL-ARTIST) with the Mann-Whitney U test to identify conditionally-essential genes (CON-ARTIST) (107). Recently, PIMMS (Pragmatic Insertional Mutation Mapping System) (108), Tn-Seq Explorer (109), and TRANSIT (110) have been published. Though they differ in the methods used for analysis and the format of the output, these methods now allow even novices to transposon sequencing to more easily analyze their data which will increase the popularity of these techniques.

Transposon studies in *S. aureus.*

Transposon mutagenesis has been used for studying the physiology of *S. aureus* for many years. In 1983, one of the first studies used Tn*551* translocated into the chromosome of a strain of MRSA to discover the first chromosomal factors essential for methicillin resistance (111). These transposon mutants turned out to map to the *femA* gene which is involved in synthesizing the pentaglycine bridge of the cell wall (112). A larger library made up of 1,012 isolated colonies was created in 1994. This library was used to expand the known number of genes that are important for β-lactam resistance. They identified 70 new insertional mutants that decrease resistance to β-lactam antibiotics (113). Studying these genes has increased our understanding not only of cell wall biosynthesis but also the requirements for β-lactam resistance in MRSA.

Over the years, transposon mutant libraries have been used study many aspects of *S. aureus* physiology. These include (but are not limited to) capsule synthesis (114), autolysis (115), β-lactamase expression (116), clumping factors (117), virulence (118), abscess formation (119), pigment production (120), and biofilm formation (121). While these studies led to a much greater understanding of *S. aureus* biology, most of these transposon libraries contained at

most ~1000 unique mutants, meaning that many genes in the genome were still untested in these experiments. Furthermore, techniques such as arbitrarily-primed PCR or inverse PCR were used to sequence the location of transposon insertion (122, 123). Such techniques are robust, but they are not high-throughput as they can be performed on only one colony at a time.

Even though genetics on *S. aureus* can be difficult, two techniques for making larger libraries were published in the early 2000's. One method uses the phage Mu and pre-assembled transposon-DNA/transposase complexes, or transpososomes. These transpososomes are assembled *in vitro* and electroporated into *S. aureus* cells, where in the presence of divalent cations, they become activated and perform the transposition reaction. This method was used to create a 10,000 colony library (124). The second method uses a mariner-based *bursa aurealis,* transposon on a plasmid, which is electroporated into cells expressing the transposase. After transposition, a high-temperature plasmid-curing step is required to cure the cells of the plasmids. The authors created a library with 10,325 unique insertions and used it to identify essential genes and new virulence factors (125).

These methods work well, and they set off an explosion of research utilizing genetic screens in *S. aureus.* Libraries and mutants created using these systems have been used to study many genes including those involved in L-form formation (126), lysostaphin resistance (127), biofilm formation (128), and fitness in blood (129). Our lab has used a library created using the plasmid-based system to identify the genes required in the absence of wall teichoic acids (95), and the same library has been used to identify genes required for fitness in infection-related ecologies such as blood and ocular fluids (90). Furthermore, the plasmid method was used to create the Nebraska Library, a transposon library created in the United States epidemic strain USA300 (130). The makers of this library have arrayed it and made it available as a resource. Unfortunately, both of these techniques suffer from a set of disadvantages. Neither protocol is simple and easy to use, mostly because of the low efficiency of electroporation into

S. aureus (131)*.* Furthermore, during the high-temperature plasmid curing step of the second method, all temperature-sensitive mutants die off (125, 132). This makes studying cell envelope and stress-response genes challenging because many of them are temperature-sensitive.

In 2011, a new method was published that solved many of the problems described above. This method uses a Φ11 phage to transduce the donor plasmid containing the transposon into the *S. aureus* strain of interest (133). This happens at a much higher efficiency than electroporation. In addition, a high-temperature plasmid-curing step becomes unnecessary because the RepC element required for replication of the donor plasmid is encoded *in trans* in the donor strain, preventing transposon plasmid replication in the recipient strain. When selecting for erythromycin resistance, which is encoded within the transposon, only those cells with a successful transposition event will survive. This system is so efficient that the authors were able to combine different transposon constructs with outward-facing promoters of different strengths when creating the library (133). Depending on the location and orientation of transposon insertion, the expression of any gene may be upregulated or downregulated (genes can be simply inactivated as well). This phage-based system was designed for predicting the target of new antibacterial compounds (133, 134). Target upregulation is a common mechanism of resistance, but may come at a fitness cost; for other antibiotics, downregulation may confer resistance (135). A major advantage of this platform is that by selecting for resistant mutants in a library with a range of expression options, it should be possible to learn something about the mechanism of action of an antibiotic and, in a best case scenario, to identify the target. The downside of this system was that it was not compatible with NGS. This means that once resistant mutants are selected on plates, the location of the transposon insertion must be identified one colony one at a time.

Looking Forward

These developments have shaped the study of *S. aureus*. Each new development builds on the last, increasing the ease with which we understand the different aspects of bacterial physiology. In the past twenty-three years, we have gone from creating small transposon libraries and identifying the location of one of the mutants, to being able to create libraries consisting of thousands of unique mutants and simultaneously identifying the location of each mutant in the library using NGS. The new "omics" era has allowed us to build enormous datasets and to begin to take a systems-level approach to understanding cellular responses to environmental stimuli. This has resulted in the creation of many tools and databases for better understanding *S. aureus*, but questions still remain. However, with the increasing threat of antibiotic resistance, these techniques will enable many functional and comparative genomics studies, which will increase our understanding of bacterial physiology, antibiotic-resistance, and host-pathogen interactions. We can use these discoveries to further our fight against these dangerous infections.

Chapter 2. Adaptation and validation of a new Tn-Seq platform in *S. aureus*

Contents of this chapter were published in BMC Genomics in 2011.

This work was done in collaboration with Tim Meredith, Leigh Matano, and Samir Moussa.

2.1 Introduction

Staphylococcus aureus is a highly adaptable pathogen that is responsible for tens of thousands of serious infections every year in the United States (136-138). Many cellular factors in *S. aureus* contribute to antibiotic resistance by mounting a robust cellular response to antibiotic-induced stress (139, 140), including the cell envelope stress response and the stringent response (141-144). A better understanding of how these cellular components work together to combat environmental stresses could lead to new strategies for more efficacious dosing of existing antibacterial agents as well as the development of novel therapeutics.

Transposon mutant libraries used in conjunction with NGS are powerful tools for probing bacterial physiology (89, 98-101, 134, 145). Creation of high-density transposon libraries in *S. aureus* has been challenging because its thick cell wall precludes high-efficiency electroporation of DNA containing the transposon, and there are no systems in *S. aureus* for transposon delivery via conjugation (124, 131). Thus, most high coverage transposon libraries in *S. aureus* typically utilize a temperature sensitive plasmid containing the transposon and require high-temperature plasmid-curing steps to remove the plasmid delivery vehicle after transposition has occurred (90, 95, 125, 130, 146, 147). During this curing step, temperature-sensitive transposon mutants may be culled, making it challenging to differentiate essential genes from those that are required for growth at elevated temperatures.

Here we report the design and application of a phage-based transposition method that is compatible with a new NGS protocol and surmounts the many challenges associated with

creating high-density transposon libraries in *S. aureus*. In fact, this method is so efficient that we were able to multiplex together transposon constructs with different regulatory elements that have the ability to over- and under-express as well as inactivate any gene in the genome to create a single highly-diverse library with more than 690,000 unique transposon mutants. We have used this library to identify genes essential for growth at 30°C and have assessed transposon mutant fitness at a set of low and high temperatures to identify temperature-sensitive and resistant mutants. Nineteen genes identified as essential in previous studies were found to be conditionally-essential, demonstrating growth inhibition at elevated temperatures. In addition, mutants in the menaquinone biosynthesis pathway were found to be significantly enriched at high temperature. Mutations in this pathway have previously been found in small colony variant (SCV) strains isolated from in vivo infections that are resistant to antibiotics, and our results suggest that heat stress is one condition that may select for these mutants. The phage-based delivery and insertion site sequencing methodology described here will facilitate comprehensive functional genomic studies of *S. aureus* physiology.

2.2 Adaptation of the transposon library for next-generation sequencing and creation of the transposon library

Adaptation of phage-based transposition for compatibility with any strain of *S. aureus*

A phage-based approach for creating transposon libraries in *S. aureus* was developed by Meredith and co-workers (133), but this method, as it was described, was not compatible with next generation sequencing. This approach uses a conditionally replicative transposon donor plasmid, which is moved by generalized phage transduction to a Himar1 transposase-expressing strain that cannot support plasmid replication (Figure 4). Following transduction, the transposase inserts the transposon into TA dinucleotide sites randomly across the genome. Whereas other transposon systems generate only null genotypes due to gene disruption, the phage-based system was designed to allow over- and under expression of genes as well as inactivation (133). This was achieved by building a set of transposon donor constructs harboring the mariner inverted terminal repeats (ITR) flanking an erythromycin resistance gene under the control of its own promoter and terminator along with an outward-facing promoter. Genes proximal to the insertion site can be upregulated or downregulated to different extents depending on the orientation of insertion and the strength of the promoter (Figure 4). This phage-based transposition system was shown to be useful for nominating antibiotic targets through gene upregulation and for identifying other cellular factors involved in antibiotic resistance through deletion/downregulation (133, 134). However, in its original format the phage-based system was limited to sequencing single colonies isolated after positive selection on agar plates. Thus, its use as a functional genomics discovery tool was limited and we sought to reengineer it to take full advantage of the power of next generation sequencing (Figure 4).
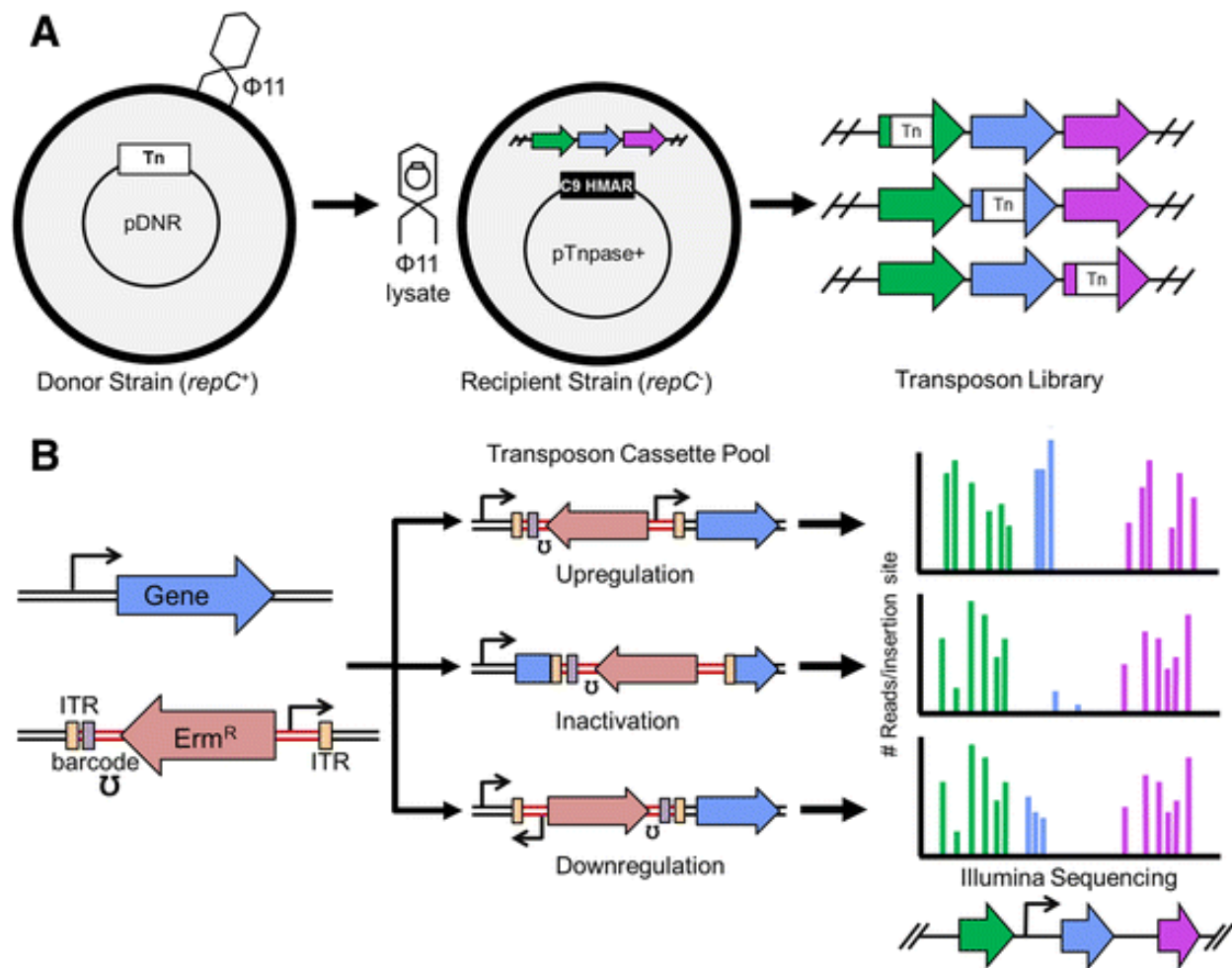
Figure 4. Strategy for using phage-based transposition to make high quality transposon libraries for NGS sequencing. (A) The transposon insertion library is made by creating a high-frequency transducing lysate of the transposon cassette that is able to replicate as a plasmid in the donor strain (*repC⁺*). The lysate is mixed with the recipient strain (*repC⁻*) carrying a temperature sensitive plasmid from which the Himar1 transposase is expressed, and erythromycin resistant transposon insertion mutants are selected. **(B)** By fitting the transposon cassette with an outward-facing promoter, genes can be up- or down-regulated, or inactivated if non-essential, in a single library pool. **Continued page 22.**

Figure 4 continued. In order to cover a wide range of gene expression levels, different promoter containing transposon constructs can be multiplexed and then identified using NGS sequencing by unique DNA barcodes (purple bar) to collect fitness-gene dose relationships by monitoring read counts. By mapping NGS reads to transposon insertion sites, and comparing the number of reads that map to each gene in the control and condition of interest, we should be able to distinguish between upregulation, downregulation, and inactivation of genes by transposon insertion.

We first modified the transposase plasmid used in the recipient strain to include *orf5*, the gene encoding the Φ11 cl-like repressor. The original phage-based transposon delivery system was used to make mutant libraries in *S. aureus* strains RN4220 and COL (133), and the transposon donor plasmids were designed with a 1 kb Φ11 homology region to stimulate efficient packaging of transposon donor plasmid DNA as concatemers by the phage. It proved necessary to integrate the Φ11 cl-like repressor (*orf5*) into the genome to prevent replication and lysis of the recipient strain by the wt Φ11 population that did not package plasmid. Because we wanted to use the phage-based transposition system in other *S. aureus* strains without having to first integrate the cl-like repressor into the genome, we moved the *orf5* gene into the transposase plasmid (pORF5) to achieve transposase expression and inhibition of Φ11 replication in a single step.

We selected HG003 as our strain background for library preparation because a high quality plasmid-delivered library has been made in the same strain and provided a reference for validating our method (90, 95). However, in contrast to RN4220 and COL, a high number of ermR colonies was observed in HG003, even in the absence of functional transposase (<1% for RN4220 versus >90% for HG003; Figure 5A). We initially speculated that phage-transposon hybrid DNA containing an *attP* site was being integrated within a recipient HG003 subpopulation

22

where the chromosomal *attB* site had become available through spontaneous excision of the resident prophage, thus leading to the high background of non-transposase catalyzed events (Figure 5B). To confirm this possibility, we introduced the pORF5 transposase plasmid into wild type RN4220, which also contains an available *attB* site. As with HG003, we now observed a high background of non-transposase catalyzed ermR colonies in RN4220 *attB+* (Figure 5A), consistent with a role for phage-mediated *att*-site specific integration in increasing background (Figure 5B). To block this pathway, we constructed a strictly virulent transducing phage (Φ11-FRT) by replacing the *attP-int* site with a FRT site from the yeast 2-μm plasmid site-specific recombination system (148, 149) (Figure 5C), thereby preventing integration. The use of Φ11-FRT decreased the background of non-transposase-catalyzed transposon integration in RN4220 as fewer ermR colonies were produced by the truncated transposase expressing strain (Figure 5A), but the background in HG003 remained unacceptably high (Figure 5A). We therefore considered a second mechanism for the production non-transposase catalyzed ermR colonies. Homologous recombination between phage-transposon hybrids carrying the ermR cassette and the resident Φ11 prophage in HG003 could also yield ermR colonies (Figure 5C). To determine whether this was occurring, we constructed a strain of HG003 where the Φ11 prophage was specifically removed using the same att::FRT exchange strategy employed to create the Φ11-FRT donor phage. The combination of this recipient strain (HG003 Φ11−) with Φ11-FRT packaged transposon donors reduced non-transposase catalyzed ermR background colonies to less than 1% (Figure 5A). With removal of the Φ11 prophage, this strategy now allows us to create high-density transposon mutant libraries with a very low background of non-transposase catalyzed transposon integration using the phage-based transposition system in any strain of *S. aureus* that is transducible by Φ11.
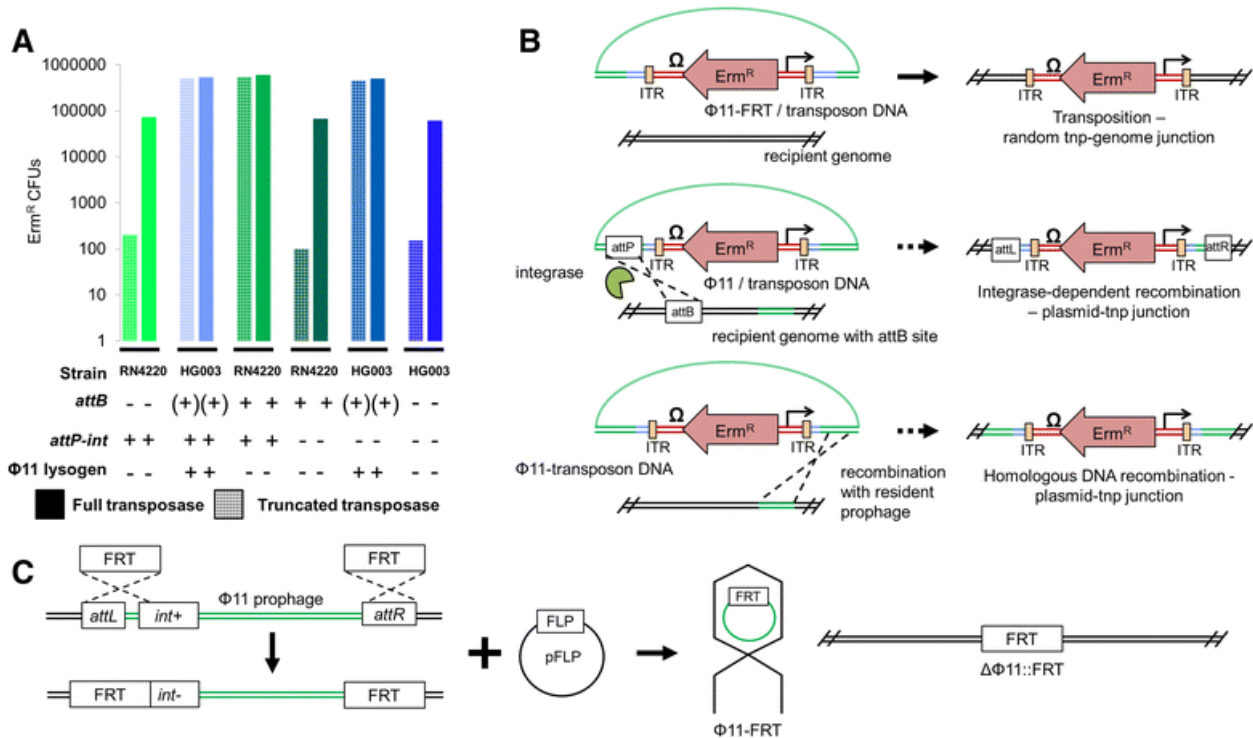
Figure 5. Elimination of non-transposase catalyzed transposon integration. (A) The ratio of ermR colonies arising from non-transposase catalyzed (hatched) to transposase dependent (solid) events was determined in the RN4220 (green) or HG003 (blue) strain backgrounds by comparing the number of colonies resulting from transduction of the transposon into the full transposase or truncated transposase expressing strains. The presence of the phage attachment site in the bacterial chromosome (*attB*), the phage attachment site in the donor lysate (*attP-int*), and a Φ11 prophage in the recipient is indicated. **(B)** Putative mechanisms for integration of the *ermR* cassette of the transposon into the recipient chromosome include transposase catalyzed (top), integrase-mediated (middle), and homologous recombination (bottom). **(C)** The integrase pathway was blocked by replacing the integrase (*int*) gene and the attL sequence with a FRT element by allelic replacement. To cure the resulting prophage, the attR site was replaced with a second FRT site and a phage donor lacking *int-attP* was isolated by introducing the FLP recombinase. In the process, a recipient HG003 strain was generated from which the Φ11 prophage was specifically cured, preventing homologous recombination.

Adaptation of the phage-based transposition system for next-generation sequencing

To adapt the transposon system for NGS, modifications to the donor plasmids were required. The initial design included the incorporation of: 1) the P7 Illumina adapter sequence within the transposon cassette to enable Illumina based NGS, 2) unique three base pair barcodes specific to each outward facing promoter element for de-multiplexing after sequencing, and 3) a MmeI site to capture the transposon-genome junction. The MmeI restriction site was embedded within one ITR of the transposon by a single base pair change, facilitating processing of transposon insertion sites by cutting non-specifically 20-base pairs downstream of its recognition site (16-base pairs downstream of the Himar1 TA dinucleotide insertion site) (99).

Early efforts to prepare and sequence our transposon libraries using the reengineered constructs were plagued by high plasmid-transposon junction read counts, despite the fact that the vast majority of ermR colonies arose via bona fide transposition events. When we used PCR to probe the transposon junctions in isolated transposon mutant colonies, we observed a small population harboring both plasmid- and genomic-transposon junctions (Figure 6A), as previously reported (133). However, when a single base pair was changed in the canonical ITR DNA sequence to create the MmeI site, the population of transposon insertion mutants containing plasmid-transposon junctions increased to over 50%. We hypothesized that the MmeI modified base in the ITR was important for recognition by the Himar1 transposase in vivo, resulting in a transposase-DNA complex that often failed to engage the initially encountered ITR and instead read through to a downstream non, contiguous ITR within the concatemer (Figure 6B) (150). The ensuing transposition event would thus capture and introduce the intervening plasmid region into the recipient genome. Therefore, we added two NotI sites to the plasmid, one immediately after the MmeI-modified ITR within the plasmid backbone and the second

immediately upstream of the P7 adapter sequence (Figure 5C). Following an added NotI

digestion step, the plasmid-transposon junctions could now be selectively removed. (Figure 6D)
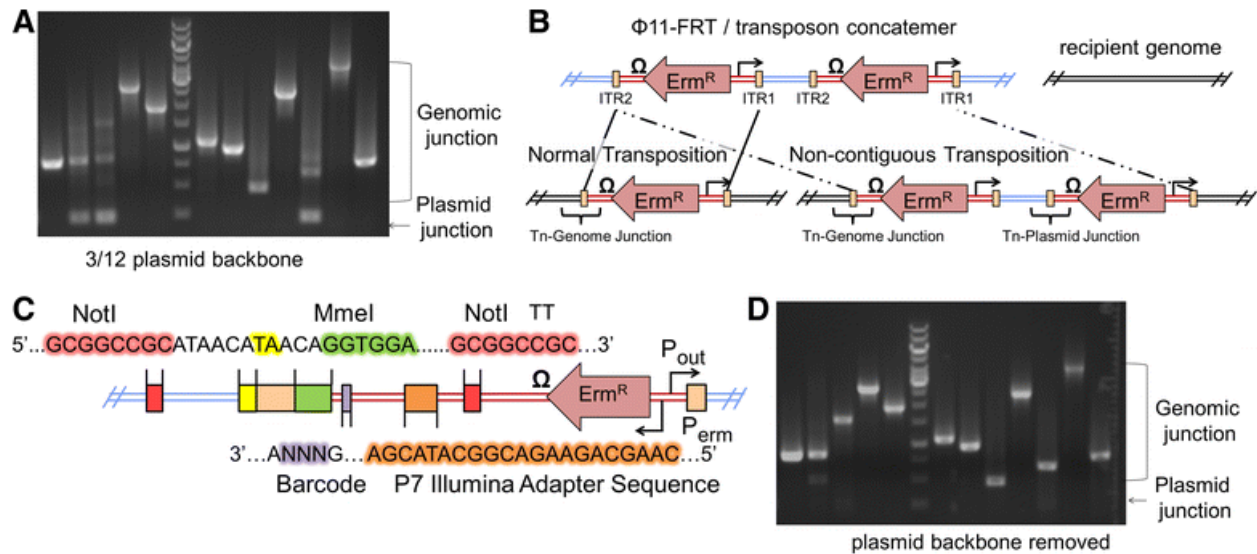
Figure 6. Reduction of transposon-plasmid junction NGS reads with flanking NotI

restriction sites. (A) Inverse PCR was used to amplify the ITR2 transposon junctions for twelve

colonies as described (33). Three out of twelve of these colonies also contained transposon-

plasmid junctions (~160 bp DNA band). This ratio increased to seven out of twelve when the

canonical ITR sequence was altered to incorporate a MmeI recognition site. Results are

representative of multiple independent experimental replicates. **(B)** The putative mechanism for

transposase catalyzed integration of transposon-plasmid junctions may involve engagement of

non-contiguous ITR repeats (dashed lines), resulting in chromosomally integrated transposon

multimers. In contrast, when both ITR sequences are optimal, contiguous ITRs are most

frequently mobilized (solid lines). **(C)** Colors are used to identify the positions of the sequences

in this drawing. To selectively remove transposon-plasmid junctions, we introduced two NotI

recognition sites into the transposon construct that flanked the MmeI modified ITR2. We also

included a P7 Illumina sequencing primer site and a unique 3-bp DNA barcode to identify the P

out promoter that faces outward from ITR1 during NGS sequencing. **(D)** After first digesting

genomic DNA with NotI, the transposon-plasmid junction content was substantially reduced.

26

Using these changes, we were able to create four different transposon constructs with outward-facing promoters of different strengths and orientations (in order of increasing strength P_{erm}, P_{pen}, P_{cap}, and P_{tuf},). In addition, we created a construct with the P_{pen} outward-facing promoter on one end and the erythromycin-resistance gene promoter (P_{erm}) on the other but no intervening transcriptional terminator (Dual), and a transposon construct with no outward facing promoter elements (Blunt) (Figure 7). Each of the six constructs is identified from the NGS data using a unique barcode located inside the transposon. The Blunt construct is most similar to transposons used in traditional transposon screening experiments, and because it does not contain any promoter elements (besides the promoter driving ErmR expression), it can only inactivate genes. The rest of the construct, which contain outward-facing P_{erm}, P_{pen}, P_{cap}, and P_{tuf} promoters, have the ability to upregulate nearby genes if these promoters are in the same orientation as that gene's native promoter. The P_{erm} construct is simply the Blunt construct without the transcriptional terminator at the end of the ErmR cassette, so the insertion orientation where it will upregulate nearby genes is opposite of the P_{pen}, P_{cap}, and P_{tuf} containing constructs. The Dual construct is unique in that, as a hybrid of the P_{erm} and P_{pen} transposon constructs, it can upregulate both upstream and downstream genes. This is useful for genes such as *tarGH.* TarGH is the wall teichoic acid transporter and both components are required for its activity. Therefore, both components would have to be upregulated to increase export of wall teichoic acids, but though *tarG* is found next to *tarH* in the genome, they face opposite directions. Only the Dual transposon mutant would be able to upregulate both *tarG* and *tarH* at once.
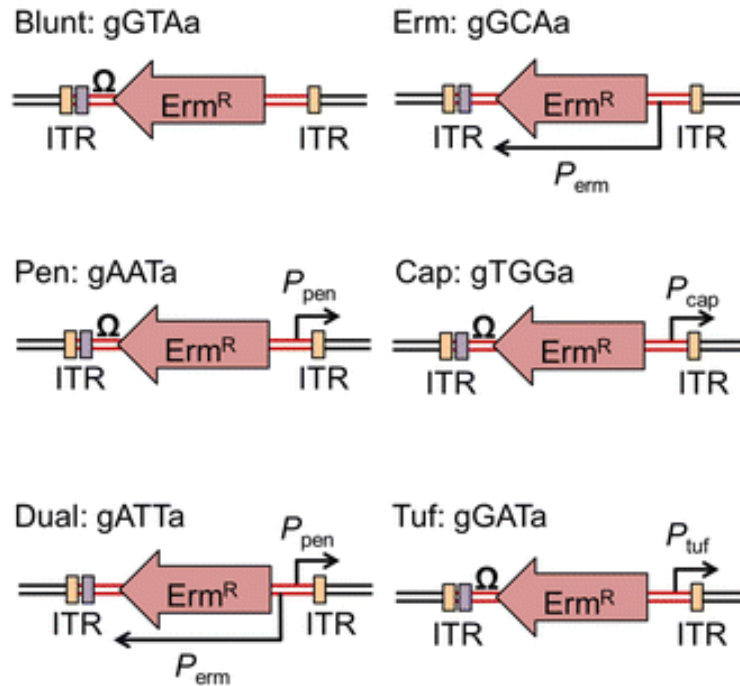
Figure 7. Six different transposon constructs were used to create the transposon library.
A schematic of the different transposon constructs, as well as the three base pair barcodes used to identify them, is shown. The Blunt construct only has the ability to inactivate genes, while the others (Erm, Pen, Cap, Dual, and Tuf) can upregulate nearby genes. The Dual transposon construct is unique in that it can upregulate genes on either side of the transposon insertion site.

Our optimized sample preparation procedure for transposon mapping is outlined in Figure 8, and a detailed version can be found in Appendix C. Briefly, genomic DNA is extracted from a pooled transposon library and then digested with NotI followed by a size-selective poly-ethylene glycol (PEG) precipitation to remove liberated plasmid-transposon junctions (151, 152). A PCR-based quality control check is performed to confirm that background due to transposon-plasmid junctions is minimal. A biotinylated dsDNA adapter containing a NotI compatible overhang is then ligated and the DNA is digested with MmeI. The transposon-genome junctions

with 2-base overhangs are bound to streptavidin dynabeads and ligated to an adapter containing the indexing barcode and priming site for the Illumina sequencing primer. Primers containing the P5 and P7 sites are then used to amplify the transposon-genome junctions bound to the streptavidin dynabeads. The fragments are run on a 2% agarose gel to confirm size, extracted from the gel, and multiplexed with other samples prior to sequencing (Figure 8). Using this protocol, we routinely reduced the amount of contaminating plasmid-transposon reads to less than 1%. To confirm the quality of the DNA insert library for NGS, we developed a quality control procedure for determining level of background due to transposon-plasmid junctions (Figure 9). This strategy, which utilizes restriction enzyme digestion followed by size-selective precipitations, can be generalized to other bacterial species and transposon library sequencing strategies to prepare transposon libraries for NGS.
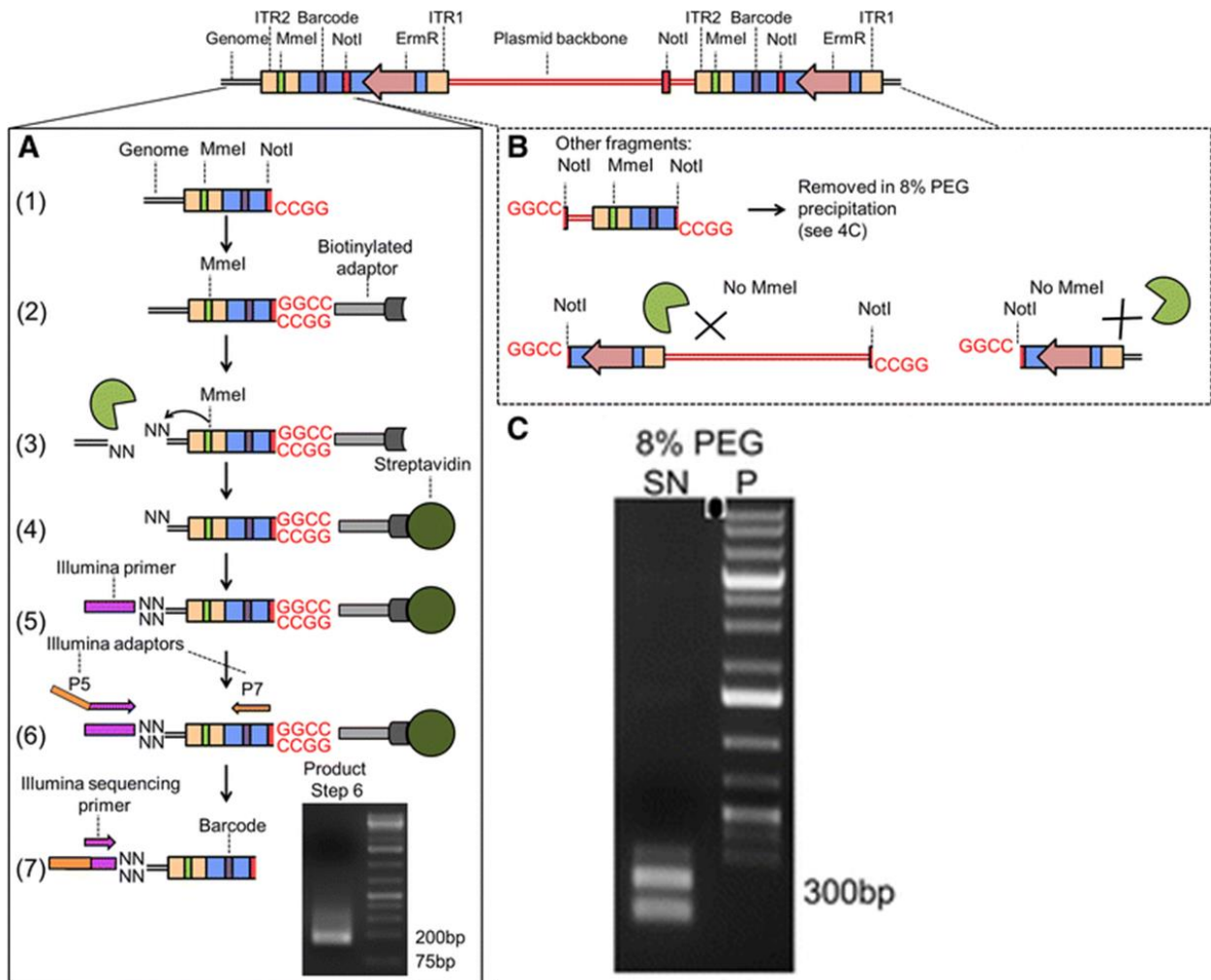
Figure 8. Protocol for the preparation of a high quality transposon DNA library for NGS.

(A) (1) Genomic DNA is isolated and digested with NotI. High molecular weight DNA is

selectively precipitated using an 8%PEG + NaCl solution. A biotinylated dsDNA adapter with

NotI overhang is ligated **(2)** before digestion with MmeI **(3)**. Biotinylated fragments are bound to

streptavidin beads **(4)**, and an Illumina sequencing primer adapter containing an indexing

barcode and MmeI compatible ends is ligated **(5)**. Primers annealing to the P7 site and the

Illumina sequencing primer adapter sequence (with a P5 site overhang) are used to PCR

amplify the transposon-genome junctions **(6)**, agarose gel purified, and submitted for Illumina

sequencing **(7). Continued page 31.**

Figure 8 continued. (B) Fragments arising from transposon-plasmid junctions are removed by size selective PEG-NaCl precipitation, while the remaining fragments lack both P7 annealing sites and MmeI sites for ligation of the Illumina sequencing primer adapter. These fragments are therefore not amplified in **(6)** of 4A. **(C)** By performing the size-selective precipitation on a 1 kb DNA ladder, we show that small 300 bp fragments of DNA are retained in the solution (SN), while larger DNA is precipitated (P).
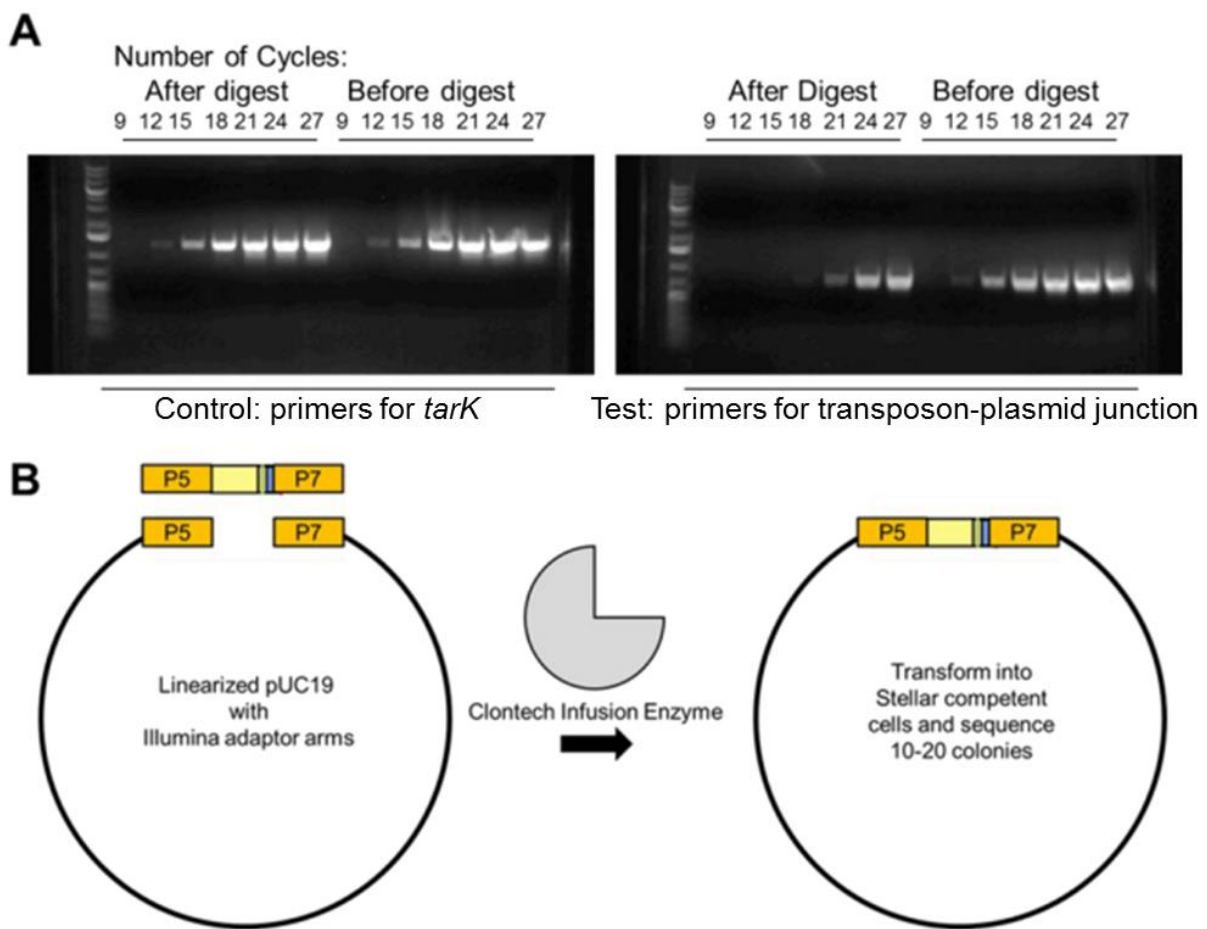
Figure 9. PCR and NGS library diversity analysis. (A) The amount of transposon-plasmid junction was quantified by removing aliquots of each PCR reaction at the indicated cycle number and analyzing by agarose gel electrophoresis. **Continued page 32.**

Figure 9 continued. The cycle at which a band was first observed for the transposon-plasmid junction had to be reduced by at least 6 cycles after NotI digestion in order for the sample to be further processed. Amplification of a genomic locus using control primers for *tarK* was used as control (top panel). **(B)** NGS library was cloned into a pUC19 vector with ends homologous to the P5 and P7 Illumina primers. Individual colonies were sequenced to confirm insert diversity before submitting for NGS.

Creation and sequencing of a transposon library

Using the strains and methods described above, we created a 2 million colony transposon library using the six different transposon donor constructs described above (Figure 7). Library cultures were grown for 12-13 generations in tryptic soy broth (TSB), harvested between an OD of 1 and 1.5, and genomic DNA was isolated. After processing samples for NGS and sequencing, the Tufts Galaxy service was used to separate the data by sample index and donor construct barcodes (153-155). The data were filtered by Illumina quality score for high quality reads and mapped with Bowtie (ver1.1) to the *S. aureus* NCTC8325 reference genome (156).

We used two biological replicates for our analyses and computationally de-multiplexed the data by transposon construct based on their unique barcode (Figure 7). For each of the donor constructs, we obtained 3 to 5 million reads that could be mapped to TA insertion sites. Depending on the constructs, insertions were identified in 105,000 to 130,000 of the ~270,000 unique TA sites in the *S. aureus* genome (Table 1). On average, $3,897,389 \pm 834,906$ reads hit $115,792 \pm 9,416$ TA sites for each of the six transposon constructs, with 36,794 TA sites in common between all six of the donor constructs, and 208,372 TA sites covered by at least one donor construct.

Table 1. Number of reads and transposon insertions from two biological replicates

Donor strain	Number of reads	Unique TA sites hit
P_{pen}	3,168,491	115,859
P_{cap}	2,938,859	105,437
P_{tuf}	4,594,924	130,003
P_{erm}	4,361,930	111,657
Dual	3,176,044	105,759
Blunt	5,126,052	126,040
Total:	**23,366,335**	**694,755**

2.3 Validation of a new Tn-Seq platform in *S. aureus* by identification of essential genes and temperature sensitivity screening

Identification of essential genes

To validate the quality of the library and assess whether insertions due to individual transposon constructs could be reliably analyzed, we used the data obtained to identify essential genes. We first compared the number of reads per gene for each transposon donor construct using principal component analysis and found that the transposon constructs with outward-facing promoters co-clustered away from the promoterless (Blunt) and weakest promoter (P_{erm}) constructs (Figure 10). This suggests that transposon insertions in the same TA site by different transposon constructs are differentially tolerated due to polar effects on downstream genes. Therefore, different numbers of reads map to each gene depending on the identity of the transposon construct. However, because transposon constructs containing a medium to strong outward-facing promoter (Dual, P_{pen}, P_{cap}, and P_{tuf}) clustered together, the number of reads mapping to each gene for each of these constructs will be more similar to each other than to Blunt or P_{erm}. Because of the similarity between the constructs with an outward-facing promoter, the data from these constructs (Dual, P_{pen}, P_{cap}, and P_{tuf}) was combined for the gene essentiality analysis.

We then identified essential genes using a recently published method, EL-ARTIST (Appendix D) (104, 107). This method uses a hidden Markov model (HMM) to categorize genes as essential, non-essential, or containing essential domains (domain essential). Because this method uses local context of the TA site as well as the number of reads in that TA site to determine essentiality, it lessens the impact of the intrinsic variability in the data and allows us to robustly ascertain the importance of every gene to cellular survival. There were 261 genes found to be essential in all three categories (Blunt, P_{erm}, and other promoter constructs). The

Blunt construct had the greatest number of essential genes (321 genes), while the P_{erm} construct had slightly fewer (295 genes), and the other promoter constructs had the least (277 genes) (Figure 10).
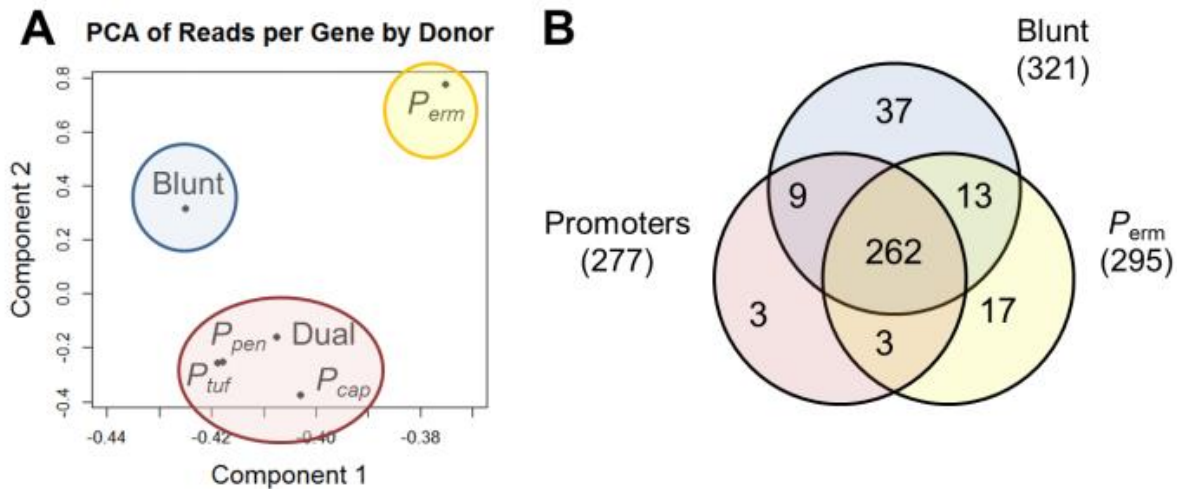
Figure 10. Transposon insertions by different transposon constructs are differentially tolerated. (A) Principal component analysis was used to compare the number of reads mapping to each gene for each of the six transposondonor constructs. The transposon constructs with outward-facing promoters clustered together, away from the Blunt and P_{erm} constructs. Data for these constructs were combined and analyzed separately. **(B)** A Venn diagram depicting the number of essential genes for each category of donor construct as calculated using EL-ARTIST is shown (107).

We were curious about the differences between the Blunt construct and constructs containing promoter elements. When a transposon inserts into the coding sequence of a gene, it not only interrupts the expression of that gene, but can also disrupt the expression of other genes due to polar effects. It is possible that some of these polar effects are abrogated by the promoters in the P_{erm} and other promoter constructs, which allow downstream genes to be expressed from the promoter in the transposon instead of the native promoter. If this is true,

then the genes found to be essential with the Blunt construct, but non-essential with the other constructs, should be next to or near other essential genes. Of 20 essential genes found using the Blunt construct that were non-essential for the P_{erm} construct and the other promoter constructs, 13 were immediately proximal to another essential gene. Another five were immediately proximal to a domain essential gene, and three of these were followed by another essential gene. Only two genes were not adjacent to an essential or domain essential gene. These data suggest that when a transposon containing an outward-facing promoter inserts into a region containing many essential and domain essential genes, the insertion may result in a less lethal phenotype than the standard transposon construct. Due to these differences and to compare our data with previous studies which did not have constructs containing promoters, we chose to use only the Blunt construct for subsequent analyses.

The Blunt construct, which is not fitted with gene expression modulating elements, is most similar to previously used transposons (90, 125, 130, 146, 147). In two biological replicates, there were ~126,000 unique insertion sites due to the Blunt construct. Using the Circos program to display the number of reads per TA site (157), we visually confirmed randomly distributed insertions throughout the length of the genome (Figure 11A). Of the 2723 coding regions of the genome that are not part of the Φ11 family lysogen, EL-ARTIST called 2212 non-essential, 190 domain essential, and 321 essential (Appendix D).

Essential genes are defined as those genes that are required for cellular growth. However, depending on the growth conditions, different sets of genes may be identified as essential, but there should be a subset of truly essential genes that are required for growth in all conditions. We expect these genes to be involved in core cellular functions such as DNA, RNA, and protein synthesis as well as the processes required for cell growth and division like peptidoglycan and membrane biosynthetic genes. A comparison with previous studies using transposon mutant analysis to identify essential genes in *S. aureus* should bring us closer to this

core set of essential genes because other studies were done using different growth media and analyzed using different methods. However, transposon mutagenesis and sequencing involves growing all the transposon mutants together in one culture, resulting in competition between mutants in the library. Therefore, non-essential genes with a fitness defect when knocked out may appear to be essential in these types of experiments because they are out-competed by other mutants in the library. We have observed this for genes such as those of the *dlt* operon, which though they can be deleted, appear essential in a Tn-Seq experiment.

We compared these results to two previous studies that identified essential genes in *S. aureus* based on transposon mutant analysis, one in the HG003 genetic background (90) and the other in SH1000 (146). Both strains are derived from the NCTC 8325 parent strain. The transposon libraries were made, grown, and analyzed using different methods, but there were 247 essential genes in common between this study and the previous study in HG003, and 211 genes in common between all three studies (Figure 11, Table 2). There were 73 genes identified as essential in both previous studies that did not meet our cutoffs for essentiality. Of these, two are no longer considered to be genes, eight were in our list of non-essential genes, and 63 were found to be domain essential.

Because the previous studies were done using transposon libraries made using a high-temperature plasmid-curing step, we checked to see if any of these genes would be identified as essential at 43°C. Of the 63 domain essential genes, nine were identified as essential at high temperatures, with nine of the remainder found to be domain essential as at 30°C. One gene, *SAOUHSC_01028*, encoding the phosphor-carrier HPr was found to be non-essential at 43°C. Transposon insertions in domain essential genes are likely to result in cells that are weak and/or have growth defects, and depending on the stringency of the method of analysis, these mutants may be mistakenly classified as essential. In addition, at the more stressful high temperatures, some of these weakened mutants may no longer be able to survive. Finally, of the eight genes

37

we found to be non-essential that were previously identified as essential, four became essential at high temperatures, and one became domain essential. The four genes that became essential at 43°C include genes known to be important for high temperatures growth such as the heat shock genes, *dnaJ* (*SAOUHSC_01682*) and *dnaK* (*SAOUHSC_01683*), as well as *sepF* (*SAOUHSC_01154*). However, overall there was substantial agreement between our essential gene list and the lists generated in previous studies despite analytical and experimental differences (different growth media, growth temperature, and/or strain background) (Appendix E). The 212 genes essential in all studies comprise a core set of essential genes required for cell survival regardless of the experimental and analytical methods used to identify them (Table 2).
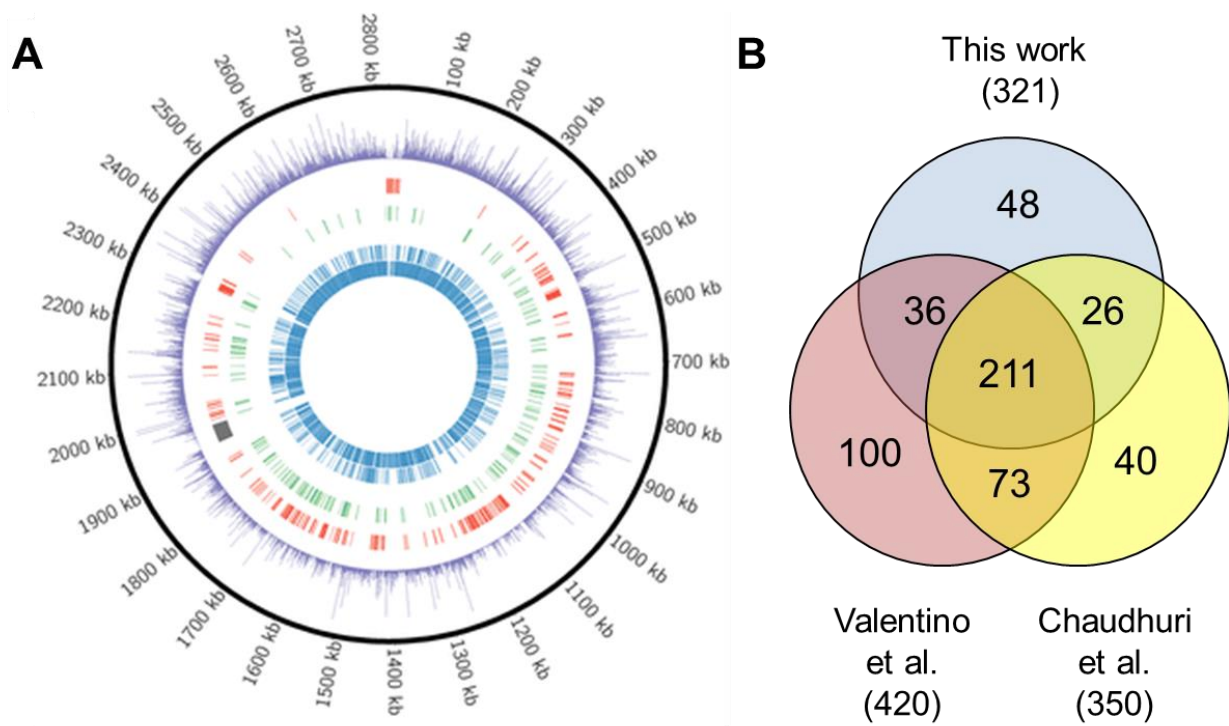
Figure 11. There are a core set of 211 essential genes in *S. aureus*. (A) The Circos program was used to map transposon insertion sites across the genome, with a histogram depicting the number of reads per TA site in purple. **Continued page 39.**

Figure 11 continued. The innermost blue rings depict the locations of non-essential genes for which a fitness cost was not observed, while the outermost red ring depicts those genes identified as essential. The middle green ring represents those genes that were found to be domain essential. **(B)** Venn diagram comparing our essential gene list to Chaudhuri et al. and Valentino et al (90, 146). There is extensive overlap between the three studies. 212 genes were classified as essential in all three works. These represent a core set of genes required for cell growth regardless of experimental and analytical variations.

Table 2. Essential genes in three studies: Core *S. aureus* essential genes

SAOUHSC_00001	SAOUHSC_00850	SAOUHSC_01287	SAOUHSC_02116
SAOUHSC_00005	SAOUHSC_00851	SAOUHSC_01351	SAOUHSC_02117
SAOUHSC_00006	SAOUHSC_00868	SAOUHSC_01373	SAOUHSC_02118
SAOUHSC_00009	SAOUHSC_00869	SAOUHSC_01374	SAOUHSC_02122
SAOUHSC_00015	SAOUHSC_00870	SAOUHSC_01424	SAOUHSC_02123
SAOUHSC_00018	SAOUHSC_00871	SAOUHSC_01434	SAOUHSC_02133
SAOUHSC_00021	SAOUHSC_00872	SAOUHSC_01467	SAOUHSC_02140
SAOUHSC_00226	SAOUHSC_00900	SAOUHSC_01470	SAOUHSC_02151
SAOUHSC_00349	SAOUHSC_00921	SAOUHSC_01473	SAOUHSC_02152
SAOUHSC_00375	SAOUHSC_00933	SAOUHSC_01474	SAOUHSC_02255
SAOUHSC_00442	SAOUHSC_00943	SAOUHSC_01492	SAOUHSC_02277
SAOUHSC_00444	SAOUHSC_00947	SAOUHSC_01496	SAOUHSC_02279
SAOUHSC_00451	SAOUHSC_00980	SAOUHSC_01501	SAOUHSC_02280
SAOUHSC_00454	SAOUHSC_01035	SAOUHSC_01592	SAOUHSC_02318
SAOUHSC_00461	SAOUHSC_01036	SAOUHSC_01598	SAOUHSC_02336
SAOUHSC_00471	SAOUHSC_01063	SAOUHSC_01599	SAOUHSC_02357
SAOUHSC_00472	SAOUHSC_01075	SAOUHSC_01623	SAOUHSC_02359
SAOUHSC_00482	SAOUHSC_01092	SAOUHSC_01624	SAOUHSC_02368
SAOUHSC_00484	SAOUHSC_01093	SAOUHSC_01662	SAOUHSC_02399
SAOUHSC_00490	SAOUHSC_01144	SAOUHSC_01663	SAOUHSC_02407
SAOUHSC_00491	SAOUHSC_01145	SAOUHSC_01666	SAOUHSC_02478
SAOUHSC_00493	SAOUHSC_01146	SAOUHSC_01672	SAOUHSC_02485
SAOUHSC_00510	SAOUHSC_01147	SAOUHSC_01690	SAOUHSC_02486
SAOUHSC_00511	SAOUHSC_01148	SAOUHSC_01697	SAOUHSC_02487
SAOUHSC_00518	SAOUHSC_01149	SAOUHSC_01700	SAOUHSC_02488
SAOUHSC_00519	SAOUHSC_01150	SAOUHSC_01701	SAOUHSC_02489
SAOUHSC_00520	SAOUHSC_01159	SAOUHSC_01722	SAOUHSC_02490
SAOUHSC_00521	SAOUHSC_01178	SAOUHSC_01727	SAOUHSC_02491
SAOUHSC_00524	SAOUHSC_01179	SAOUHSC_01738	SAOUHSC_02492
SAOUHSC_00525	SAOUHSC_01183	SAOUHSC_01741	SAOUHSC_02493
SAOUHSC_00527	SAOUHSC_01189	SAOUHSC_01746	SAOUHSC_02494
SAOUHSC_00528	SAOUHSC_01190	SAOUHSC_01753	SAOUHSC_02495
SAOUHSC_00529	SAOUHSC_01197	SAOUHSC_01756	SAOUHSC_02496
SAOUHSC_00530	SAOUHSC_01198	SAOUHSC_01757	SAOUHSC_02498
SAOUHSC_00574	SAOUHSC_01199	SAOUHSC_01766	SAOUHSC_02499
SAOUHSC_00578	SAOUHSC_01201	SAOUHSC_01767	SAOUHSC_02500
SAOUHSC_00579	SAOUHSC_01205	SAOUHSC_01784	SAOUHSC_02501
SAOUHSC_00611	SAOUHSC_01207	SAOUHSC_01785	SAOUHSC_02502
SAOUHSC_00641	SAOUHSC_01209	SAOUHSC_01786	SAOUHSC_02503
SAOUHSC_00643	SAOUHSC_01222	SAOUHSC_01787	SAOUHSC_02504
SAOUHSC_00742	SAOUHSC_01234	SAOUHSC_01791	SAOUHSC_02505
SAOUHSC_00743	SAOUHSC_01235	SAOUHSC_01792	SAOUHSC_02506
SAOUHSC_00752	SAOUHSC_01236	SAOUHSC_01807	SAOUHSC_02507
SAOUHSC_00769	SAOUHSC_01238	SAOUHSC_01809	SAOUHSC_02508
SAOUHSC_00771	SAOUHSC_01240	SAOUHSC_01811	SAOUHSC_02509
SAOUHSC_00785	SAOUHSC_01241	SAOUHSC_01829	SAOUHSC_02510
SAOUHSC_00795	SAOUHSC_01243	SAOUHSC_01839	SAOUHSC_02511
SAOUHSC_00796	SAOUHSC_01244	SAOUHSC_01856	SAOUHSC_02512
SAOUHSC_00797	SAOUHSC_01245	SAOUHSC_01871	SAOUHSC_02527
SAOUHSC_00799	SAOUHSC_01246	SAOUHSC_01875	SAOUHSC_02623
SAOUHSC_00804	SAOUHSC_01249	SAOUHSC_02106	SAOUHSC_03054
SAOUHSC_00848	SAOUHSC_01252	SAOUHSC_02107	SAOUHSC_03055
SAOUHSC_00849	SAOUHSC_01260	SAOUHSC_02114	

Identification of genes important for growth at different temperatures

Because the phage-based method for transposon delivery does not involve a high temperature plasmid curing step, mutant libraries retain insertions in temperature sensitive genes. Hence, after validating the transposon library grown at 30°C, we identified transposon insertions that confer temperature-sensitive and resistant phenotypes by comparing mutant fitness at 16°C, 23°C, 37°C, and 43°C to that at the reference temperature 30°C. At each temperature the library was grown for 12-13 generations and the number of reads due to insertion of the Blunt transposon in each gene was compared to the number of reads obtained after outgrowth at 30°C. Two biological replicates were carried out for each condition and the data were analyzed using the Mann-Whitney U test, correcting for multiple hypothesis testing with the Benjamini-Hochberg procedure (95). Genes showing a greater than five-fold change (increase or decrease) in the number of reads at the test temperature compared to the 30°C control were considered significant if the corrected p-value was less than 0.05. A full list of affected genes can be found in Appendix F.

At 16°C and 23°C, only one gene, *SAOUHSC_01857*, contained a significantly different number of reads due to transposon insertions compared to the 30°C control, with the number of reads being enriched six-fold and five-fold, respectively. At 37°C, reads due to transposon insertions were found to be enriched in five genes involved in branched chain amino acid degradation, *ilvE* (*SAOUHSC_00536), SAOUHSC_01611*, *bkdA2* (*SAOUHSC_01612*), *bkdA1* (*SAOUHSC_01613*), and *SAOUHSC_01614* (*SAOUHSC_01611*, *bkdA1*, *bkdA2*, and *ilvE* meet significance cutoffs) (Figure 12).
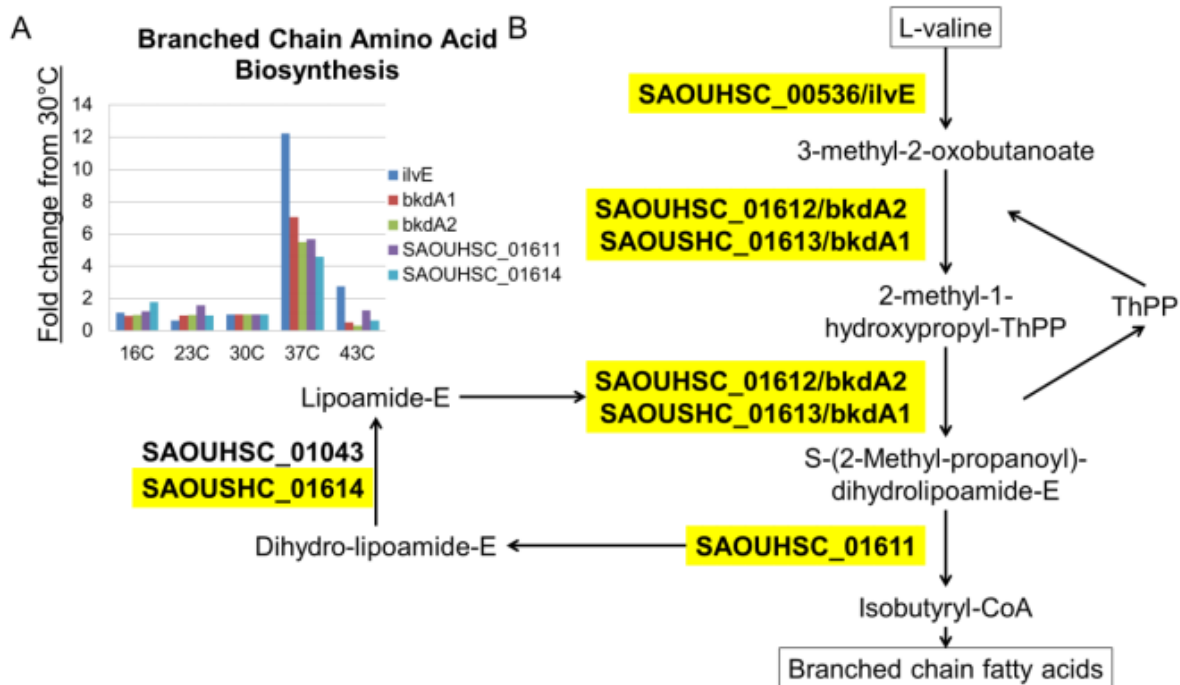
Figure 12. Enrichment of transposon insertions in the branched chain amino acid degradation pathway. (A) The number of reads in each gene per 5 million reads was normalized to the read count at 30°C and plotted for each temperature. **(B)** Four genes in the essential branched chain fatty acid pathway were found to have a statistically-significant increase in the number of reads and *SAOUHSC_01614* was found to have a non-statistically-significant increase in number of reads at 37°C, suggesting they have less impact on fitness at this temperature.

Branched chain fatty acids are built from branched chain acyl-CoA primers, which in turn are derived from the degradation of branched chain amino acids (158). Branched chain fatty acids increase membrane fluidity upon incorporation into phospholipid bilayers (159). In *Bacillus subtilis*, enzymes that catalyze the degradation of branched-chain amino acids are induced during cold shock and are speculated to increase membrane fluidity by promoting the incorporation of branched chain fatty acids into membrane lipids (160). In our data there is a

bias against transposon insertions in the branched chain amino acid degradation pathway at temperatures lower than 37°C, in agreement with studies in *S. aureus* and *Listeria monocytogenes* implicating branched chain fatty acids in growth at low temperatures (159, 161). Two genes, *SAOUHSC_01154*, encoding SepF, and the hypothetical protein likely co-transcribed with it, *SAOUHSC_01153*, proved to have significantly fewer reads at 37°C. SepF is a protein of unknown function thought to be involved in cell division, and was previously identified as essential in *S. aureus* (90, 125, 130, 146, 147), but these studies show it is essential for survival only at temperatures greater than 30°C.

A significant number of genes were found to be affected by transposon insertion when grown at 43°C, with 42 being enriched and 77 being depleted. Because this method for creating transposon libraries does not involve a high-temperature plasmid curing step, we expected to retain many temperature-sensitive mutants. We used two methods for identifying temperature-sensitive mutants represented in the library. The first was the Mann-Whitney U analysis used for every temperature. For the second analysis, we used the 43°C data and the essential gene analysis described in the previous section to identify genes essential at 43°C that were not identified as essential at 30°C. We confirmed that the number of reads mapping to these genes had decreased at 43°C from 30°C. We compared these genes to other essential genes lists (90, 125, 130, 146, 147), and identified 19 genes that had been annotated as essential in at least two other transposon library analyses, but were only temperature-sensitive with our method (Table 3). To further analyze the temperature sensitive gene subset, we classified them according to cellular function (Figure 13A).

Table 3. Genes with a growth defect at 43°C that were previously identified as essential

Category	Gene Locus	Name
Cell envelope	SAOUHSC_01154	*sepF* [90,125,130,146,147]
	SAOUHSC_01359	*mprF** [90,146,147]
	SAOUHSC_01627	Putative lipoprotein [90,125,130,146,147]
	SAOUHSC_01739	*lytH* [146,147]
	SAOUHSC_01759	*mreC* [90,147]
	SAOUHSC_02319	*rodA* [90,147]
	SAOUHSC_01827	*ezrA* [90,125,130,146,147]
	SAOUHSC_02337	*murA* [146,147]
	SAOUHSC_02571	*ssaA* [90,125,130,146]
	SAOUHSC_01361	*lcpA** [90,147]
Protein folding	SAOUHSC_01682	*dnaJ* [90,125,130,146,147]
	SAOUHSC_01683	*dnaK* [90,125,130,146,147]
	SAOUHSC_01684	*grpE* [90,125,130,146,147]
Other	SAOUHSC_00803	Ribonuclease R [90,147]
	SAOUHSC_00015	DhH subfamily protein [90,125,130,146]
	SAOUHSC_00892	General stress protein 13 [125,130,147]
	SAOUHSC_01751	*ruvA* [125,130,146,147]
Unknown	SAOUHSC_00760	Hypothetical protein [146,147]
	SAOUHSC_00788	Hypothetical protein [90,146,147]

* *mprF* also annotated as *fmtC,* and *lcpA* also annotated as *msrR*

Because protection from temperature stress relies on the coordinated response of many different pathways regulated by signaling systems and alternative transcription factors, we were not surprised to find that 11 regulatory genes were significantly depleted at 43°C, including *yycH* (*SAOUHSC_00022*) and *yycI* (*SAOUHSC_00023*), which were identified as essential in another transposon library (90). The *yycH* and *yycI* genes negatively regulate the two component system, *walKR*, which is a major regulator of peptidoglycan metabolism that controls autolytic activity (162, 163). Deletions of *yycH* and *yycI* result in upregulation of *walKR*, which induces

cell wall defects (162, 164). As *walKR* positively regulates autolytic activity (165), increased autolysis of peptidoglycan could explain how derepression of *walKR* produces a temperature-sensitive phenotype.

We also found, as expected, that many genes implicated in the heat shock response were required for survival at 43°C. These included the chaperones *dnaK* (*SAOUHSC_01683*) and *dnaJ* (*SAOUHSC_01682*) (166), the transcriptional regulator *hrcA* (*SAOUHSC_01685*) (167), the protease *clpC* (*SAOUHSC_00505*) (168), and *mcsB (SAOUHSC_00504*; also known as *yakI*), an arginine phosphotransferase that negatively regulates the stress response repressor *ctsR* (169, 170). GrpE (*SAOUHSC_01684*), which acts in a complex with DnaK and DnaJ, narrowly missed our cutoffs (Figure 13B) (171).
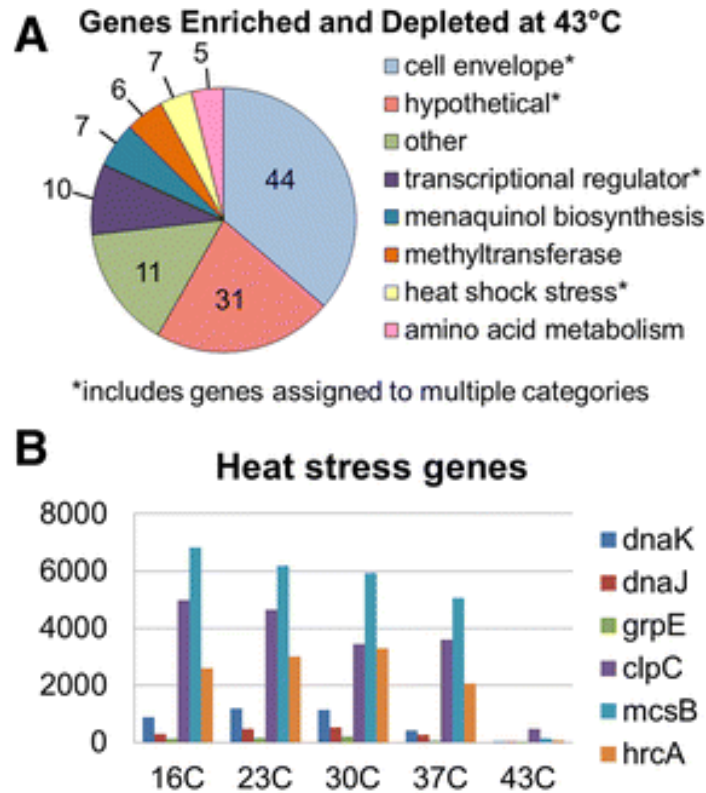
Figure 13. Genes that influence fitness at high temperature. (A) Genes contributing to fitness at 43°C were classified by cellular function. Genes associated with the cell envelope constituted the largest fraction of the hits. **(B)** Six genes known to be associated with the heat-shock response were found to be temperature-sensitive. The number of reads mapping to these genes at different temperatures normalized to 5 million reads is shown.

The largest category of genes with significant changes in number of reads when grown at 43°C included those involved in cell envelope related processes (44 genes) (Figure 13A). This category includes genes involved in peptidoglycan (PG) biosynthesis, cell shape/division, membrane lipid composition, transcriptional regulation of cell-envelope genes, and predicted membrane proteins. Transposon insertions into biosynthetic genes directly involved in PG biosynthesis were notably depleted, including *pbp3* (*SAOUHSC_01652*) and *pbp4* (*SAOUHSC_00646*), which encode cell wall transpeptidases (172-174), one of two alleles of

murA (*SAOUHSC_02337*) involved in PG precursor biosynthesis (175), *alr1* (*SAOUHSC_02305*) which encodes one of two alleles of alanine racemase (176), and *SAOUHSC_01739*, which encodes an amiC-like peptidoglycan amidase (177). Insertions in several genes involved in cell shape/cell division were also affected at 43°C. For example, reads due to transposon insertion were depleted for *rodA* (*SAOUHSC_02319*), *mreC* (*SAOUHSC_01759*), and *mreD* (*SAOUHSC_01758*), which encode scaffolding proteins that coordinate PG biosynthesis and influence cell shape (178-181), and also for *ezrA* (*SAOUHSC_01827*), which is thought to coordinate Z-ring functions with PG synthesis (182), *ftsH* (*SAOUHSC_00486*) (183), *sepF* (*SAOUHSC_01154*) (184), *gpsB* (*SAOUHSC_01462*) (185), and *SAOUHSC_01857*. *SAOUHSC_01857* encodes a 1200 amino acid FtsK-like protein suggested to be involved in chromosome localization (186). As noted above, insertions in this gene conferred a growth advantage at 16°C and 23°C, but are deleterious at high temperature. This gene was identified in other studies as essential in *S. aureus* (90, 125, 130, 146, 147). The putative cell wall teichoic acid ligases, *lcpA* (*SAOUHSC_01361*) and *lcpB* (*SAOUHSC_00997*), were also depleted at 43°C (25-fold and 5-fold, respectively), consistent with the vital role of wall teichoic acids in orchestrating PG assembly and maintaining envelope integrity (187-189). The integral membrane protein MprF (*SAOUHSC_01359*), which attaches lysine to phosphatidylglycerol, was also important for fitness at high temperatures (190).

In addition to the mutants that were depleted at 43°C, we identified a number of processes for which disruption resulted in a significant growth advantage compared to growth at 30°C. Pathway enrichment analysis using BioCyc (191) showed that the number of reads due to transposon insertion were significantly enriched in seven genes in the aromatic amino acid and menaquinol biosynthetic pathways (*SAOUHSC_01481, aroB: SAOUHSC_01482, aroF: SAOUHSC_01483, menF: SAOUHSC_00982, menD: SAOUHSC_00983, aroE: SAOUHSC_01699*, and *aroA: SAOUHSC_01852*) (Figure 14). In addition to producing

phenylalanine and tyrosine, the aromatic amino acid pathway provides precursors (chorismate)

for menaquinone biosynthesis (192). Menaquinones are isoprenylated electron transport chain

cofactors embedded in the membrane that are necessary for oxidative phosphorylation (193).

Insertions in *ispA* (*SAOUHSC_01618*), which encodes a putative geranyltransferase, were also

enriched at 43°C. Geranyltransferases initiate the condensation of isoprenoid units into longer

chains used in the synthesis of menaquinones, and loss of *ispA* likely decreases the pool of

menaquinones. We also identified insertions in *hemY* (*SAOUHSC_01460*), a gene involved in

the production of protoheme (194). Loss of these factors is thought to shift *S. aureus* growth

towards an anaerobic metabolic regime, which markedly impacts *S. aureus* membrane

physiology and generates small colony variants (SCV) (195, 196). The observed growth

advantage could also be related to more efficient anaerobic catabolism at high temperature,

and/or decreased oxidative stress. Five subunit genes in the F-type ATPase involved in electron

transport were found to be essential at 43°C (Figure 12B): α-subunit (*SAOUHSC_02345*), β-

subunit (*SAOUHSC_02343*), γ-subunit (*SAOUHSC_02341*), A subunit (*SAOUHSC_02350*), and

B subunit (*SAOUHSC_02347*) (all were significantly depleted except the β-subunit), suggesting

the electron transport system (ETS) mutations can have opposing effects on fitness at elevated

temperatures depending on which step in the ETS is blocked.
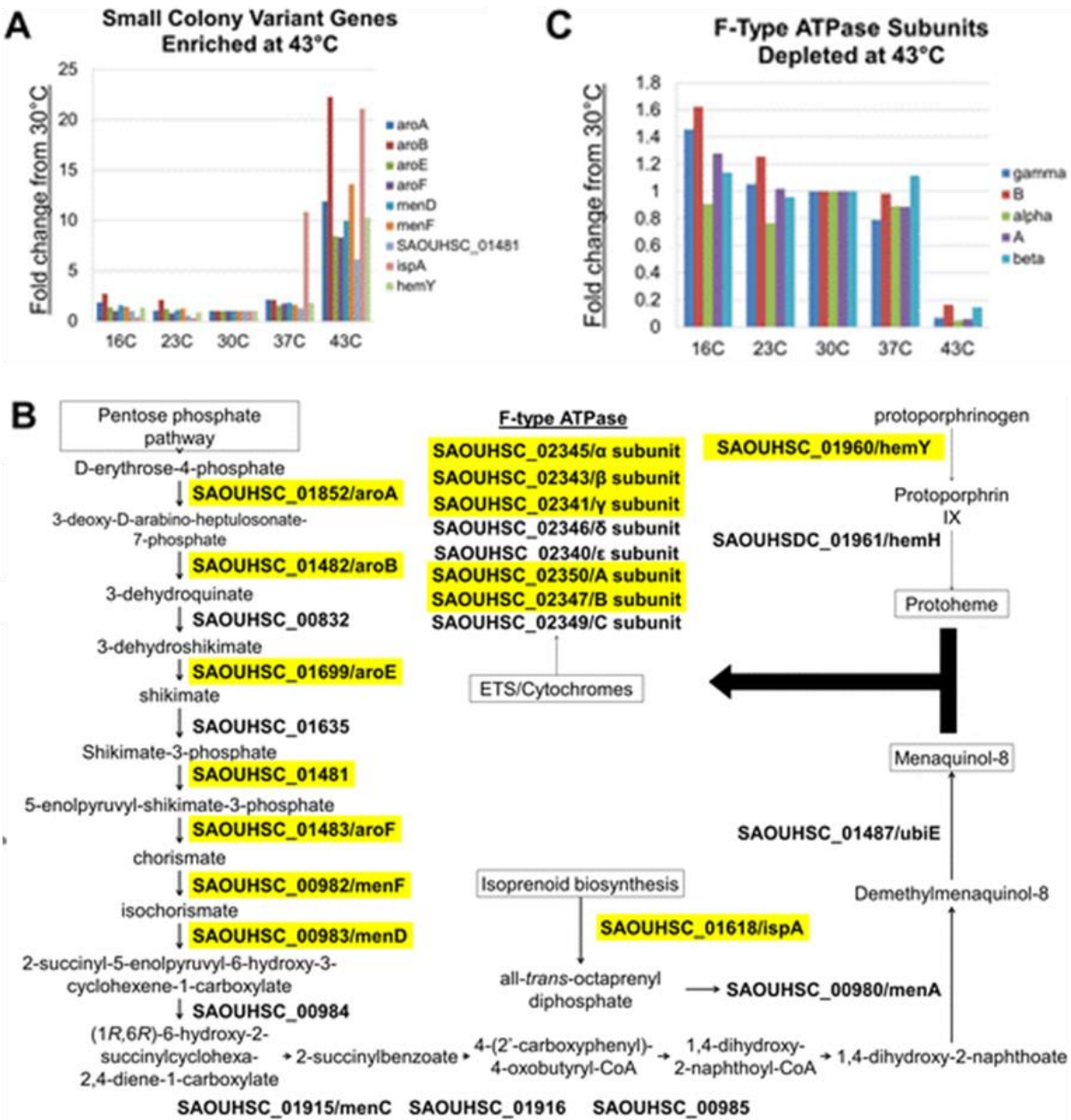
Figure 14. The electron transport system influences sensitivity to high temperatures. (A)
Nine genes were identified as growth advantaged at 43°C in comparison to 30°C. The number
of reads per gene normalized to 5 million reads is shown for each temperature. **(B)** Pathway
analysis revealed that this subset of genes (highlighted in yellow) is involved in the biosynthesis
of components within the electron transport system, including protoheme and menaquinones.

Figure 14 Continued. (C) Blocking the electron transport system at the F-type ATPase level, however, decreased fitness at 43°C (genes highlighted in yellow). The number of reads per gene normalized to 5 million reads is shown for each temperature.

Among cell envelope-related genes, reads due to transposon insertion were notably enriched in *lytD* (*SAOUHSC_01895*, also known as *sagB*), encoding β-N-acetylglucosaminidase (196, 197), and *lyrA* (*SAOUHSC_02611*), a polytopic membrane protein. Disruptions in *lyrA* were previously shown to increase lysostaphin resistance and are lethal when wall teichoic acid biosynthesis is blocked (95, 127), and both these phenotypes are consistent with an important role for *lyrA* in cell envelope biogenesis.

To confirm the results of our automated transposon insertion analysis pipeline, we constructed and measured the growth rates of null mutants of *lyrA* (significantly overrepresented at 43°C) and *mprF* (significantly depleted at 43°C) (Figure 15A and B). While there were no differences in growth of the mutants compared to wild type at 30°C, growth related phenotypes became apparent at 43°C after 5 to 6 generations of outgrowth. The wild type strain grew more slowly at 43°C than at 30°C (doubling times were 65 min and 37 min, respectively), the *ΔmprF* strain ceased dividing altogether, and the doubling time of the *ΔlyrA* mutant decreased to 34 minutes at 43°C from 41 minutes at 30°C (Figure 15C). A faster growth rate than wild type could signify that in the absence of *lyrA,* an important mechanism for regulation of growth rate during conditions of stress such as high heat has been disrupted. On the other hand, an increase in OD, does not necessarily mean that all the cells are alive and growing. There may be many dead cells in this condition, contributing to a higher OD, but not actually corresponding to an increase in fitness at 43°C. These results confirm the changes in fitness that were deduced from the changes in read number due to transposon insertion in *lyrA* and *mprF* at 43°C.
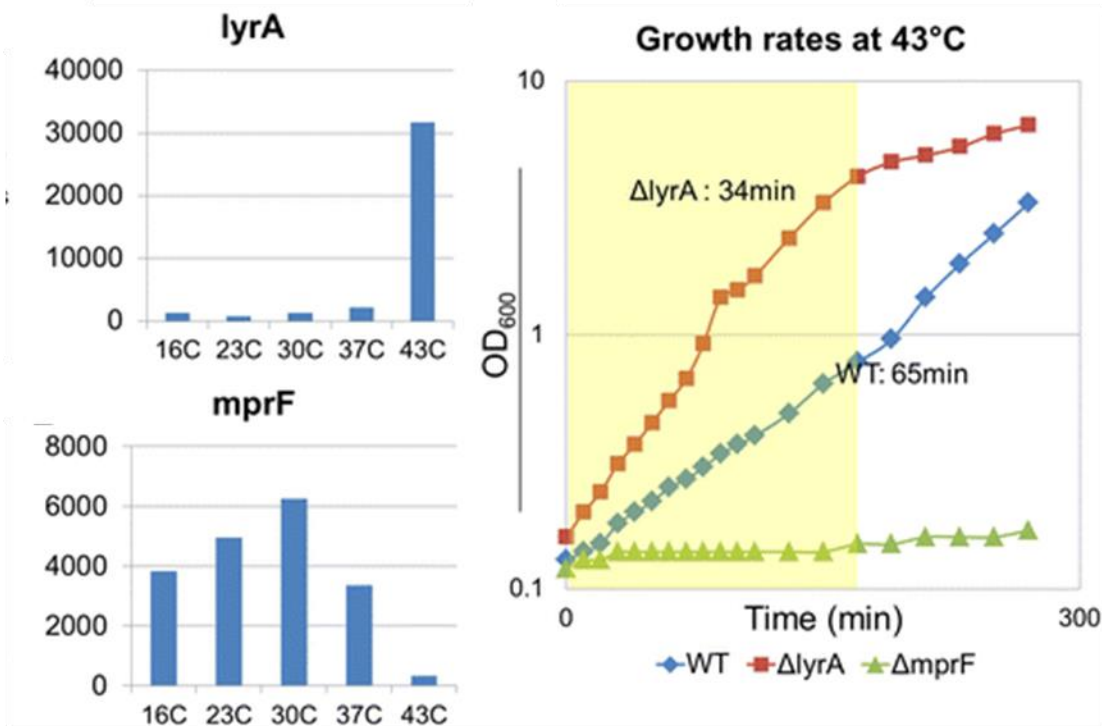
Figure 15. Confirmation that changes in read number correspond to changes in fitness.

The number of reads in each gene normalized to 5 million sample reads is shown at the various

temperatures tested. Loss of *lyrA* and *mprF* was found to have opposite phenotypes at 43°C,

with an increase **(A)** and decrease **(B)**, respectively, in number of reads mapping to each of

these genes. **(C)** To validate these phenotypes, null mutants *in S. aureus* HG003 were grown to

mid-log phase, diluted, and grown at 43°C. While *ΔmprF* did not grow, confirming temperature

sensitivity, the *ΔlyrA* strain grew at nearly twice the rate of WT. Doubling times for WT and the

ΔlyrA strains are shown.

2.4 Discussion

Adaptation of the transposon library for NGS

We have adapted a phage-based transposition method to reliably create large and diverse transposon libraries that can be analyzed using high-throughput Illumina sequencing. The high background due to non-transposase-catalyzed recombination in the HG003 strain was reduced to negligible levels by removing the Φ11 family prophage and using a strictly virulent transposon donor phage (Φ11-FRT) for transduction. Other strains that harbor Φ11 family prophages, such as *S. aureus* Newman (198), may have a similar background problem that makes removing the Φ11 family prophage necessary for working in these genetic backgrounds. Fortunately, many commonly used *S. aureus* strains, including RN4220, COL, USA300, and MW2, do not harbor highly similar prophages and need not be reengineered. Indeed, we have successfully constructed ultra-high density transposon libraries in several other *S. aureus* strains, including the USA300 and MW2 backgrounds, using the same tools described here (See Chapter 3).

To realize the full potential of the phage-based delivery system for transposon library generation, we optimized both the donor constructs and the NGS sample preparation procedure. Now, transposon libraries can easily and reproducibly be created in *S. aureus* and sequenced using NGS. Moreover, because transposition efficiency is high and the NGS library preparation protocol is robust, we are able to multiplex several different transposon constructs in the same library. By adding a DNA barcode that specifies the gene expression regulatory element in the transposon construct, the roles of over- and under- expression of genes as well as inactivation can be assessed under a given condition during data analysis. For functional genomics and systems level analyses, the ability to capture over- and under-expression

genotypes is particularly important. For the first time we will be able to probe the relationship of essential genes to a given cellular process through over/under expression of essential genes.

Gene essentiality at a set of temperatures

The power of the redesigned phage-based transposition system was demonstrated through the preparation of a transposon mutant library containing ~694,000 independent insertions in the *S. aureus* HG003 strain background. We identified essential genes at 30°C, and found good agreement between our essential gene list and two other lists published previously, although 63 genes previously annotated as essential were classified by us as domain essential (90, 146). Taking advantage of the fact that the phage-based transposon delivery method does not involve a high temperature plasmid curing step, we also carried out a study of the genetic factors involved in withstanding temperature stress. Many genes responsive to heat- or cold-shock have been identified in *S. aureus* based on transcriptome data (199), but a systematic and comprehensive analysis of the fitness of temperature-sensitive or resistant mutants has not been reported. We assessed the importance of each gene at five different temperatures based on enrichment or depletion of corresponding NGS reads. While reads per gene were largely unchanged compared to the 30°C control between 16°C, 23°C, and 37°C, reads for 119 genes were significantly affected at 43°C.

In addition to genes previously linked to heat shock and stress responses, we also found that many genes involved in cell envelope processes contribute to fitness at elevated temperature. Some genes in which reads were depleted at 43°C, such as *pbp3* and *pbp4*, directly participate in crosslinking PG (172, 173), but others play less well understood roles in PG synthesis. For example, our studies have implicated *rodA*, *mreC*, and *mreD* in withstanding temperature stress, as reads in all three genes were depleted by ~50-fold at 43°C. In rod-shaped organisms, different biosynthetic machines are dedicated to septal and side wall PG synthesis. RodA, MreC, and MreD are components of the machine that makes side wall PG,

while FtsZ, FtsL, and Div1B interact with FtsW as part of the PG biosynthetic machinery at the septum (178, 200-202). *S. aureus* is a coccoid organism and is not known to have distinct PG biosynthetic machines, and yet it contains genetic elements suggestive of both a primary PG biosynthetic machine and a secondary machine (203-205). In contrast to RodA/MreC/MreD, FtsZ, FtsL, Div1B, and FtsW were deemed essential from our analysis and are likely part of the predominant PG biosynthetic apparatus. RodA/MreC/MreD may be part of a secondary machine that takes on an important cellular role under environmental stress conditions.

We also identified *mprF* as important for survival at 43°C (16.7 fold depletion). MprF attaches lysine to phosphatidylglycerol in the cell membrane and is involved in cationic antimicrobial peptide resistance (190, 206). Mutants in *mprF* have not previously been shown to be temperature sensitive. Therefore, we constructed an *mprF* deletion strain and confirmed that it exhibited a pronounced growth defect at high temperature (Figure 11) as it ceased growing completely by mid-log phase. When these cells were plated and grown at 30°C, colonies did not grow, confirming that without the *mprF* gene, high temperatures are lethal.

In contrast to *mprF*, loss of another cell envelope gene, *lyrA*, conferred a growth advantage. Reads due to transposon insertion in this gene were substantially enriched at 43°C compared to 30°C. The growth advantage of a *ΔlyrA* strain compared to wildtype *S. aureus* was confirmed (Figure 11). Deletion of *lyrA* was previously shown to impart resistance to lysostaphin, an oligopeptide protease that cleaves Gly-Gly bonds in crosslinked PG (127, 207, 208). Recent work has implicated LyrA in display of cell surface proteins (209, 210), but how its deletion confers increased growth rates at higher temperature is not clear. As an integral membrane protein, LyrA may regulate multiple aspects of cell envelope structure. Perhaps through the loss of interactions with other protein partners, the absence of LyrA induces a pleiotropic cell envelope phenotype capable of withstanding a variety of cell envelope stressors.

Small colony variants withstand heat stress

Transposon insertion mutants with reads mapping to genes that are part of the aromatic amino acid and menaquinol biosynthetic pathways were significantly enriched at 43°C (Figure 12). The small colony variant (SCV) phenotype of menaquinone (vitamin K2) mutants in *S. aureus* has been well studied (211-218). The loss of electron shuttling redox cofactors in the electron transport chain induces a characteristic colony morphology, noted for its slow growth, lack of pigmentation, and pinpoint colony size (216, 218-222). This reduction in flux through the electron transport chain decreases ATP pools by diminishing respiration, which shifts global metabolism towards fermentation (220, 223). Emergence of SCVs is closely associated with persistent infections (224, 225), and SCVs are particularly resistant to many clinically used antibiotics (226, 227). Interestingly, our data revealed a marked growth advantage for menaquinone biosynthetic mutants at 43°C, including *menD, menC*, and *menF*. Mutations in these same genes have been associated with SCV formation in *S. aureus* clinical isolates (211, 216). Elevated temperature clearly acts as a selective pressure that favors the emergence and propagation of SCVs in vitro. Recently, strains with increased resistance to vancomycin were found to emerge in an in vivo mouse model independent of vancomycin treatment (228). Whether the host pyrogenic response induces SCV selection in vivo, however, remains an outstanding question.

Conclusion

In conclusion, we have developed a phage-based method for reproducibly creating high quality, ultra-high density transposon libraries in any *S. aureus* strain transducible by Φ11. We have optimized this platform to selectively remove transposon-plasmid junctions, and our solutions for reducing background due to these junctions are likely to be useful for other *S. aureus* strains, other bacterial species, and other transposon systems. By multiplexing bar-coded donor constructs, we showed that we can make massive transposon mutant libraries and

identify insertions due to each type of donor. In this study, we used the data for the Blunt transposon construct to identify genes important for *S. aureus* survival at high temperatures. Furthermore, we found that SCV mutants are selected for under conditions of heat stress. This system, including the gene expression modulating transposon cassettes, will be useful for future functional genomics analyses aimed at establishing novel strategies for the development of antibacterial agents and garnering insights into the biology of *S. aureus*.

Chapter 3. Using β-lactams to predict gene functions in MRSA

This work was done in collaboration with Tim Meredith and Samir Moussa.

3.1 Introduction

Both methicillin-resistance (MRSA) and methicillin-sensitive (MSSA) strains of *S. aureus* can be highly-virulent and cause life-threatening infections, but it is MRSA that causes the vast majority of lethal infections (136, 138). MRSA strains are defined as those that have acquired the SCCmec mobile genetic element encoding penicillin binding protein 2a (PBP2a), a transpeptidase with a low affinity for β-lactam antibiotics (except for ceftaroline and ceftobiprole) (229, 230). PBP2a can perform the peptidoglycan transpeptidation reaction, which crosslinks the glycan strands of peptidoglycan, when the other PBPs are inactivated by a β-lactam antibiotic (230, 231).

MRSA encodes four other PBPs besides PBP2a, three high molecular weight PBPs (PBP1, PBP2, and PBP3) and one low molecular weight PBP (PBP4) (232). However, only PBP1 and PBP2 are essential for cell survival (233). PBP2 is the only class A PBP, with both a transpeptidase and glycosyltransferase domain. In the presence of β-lactam antibiotics, it is thought that the glycosyltransferase domain of PBP2 cooperates with PBP2a to synthesize peptidoglycan (231, 234-236). PBP1 and PBP3 are class B PBPs with a transpeptidase domain and another domain of unknown function (232). PBP1 is thought to have roles in cell division at the cell septum that are independent of its transpeptidase activity, and it has been suggested that it may mediate important protein-protein interactions (237, 238). Its function appears to be especially important for survival with daptomycin treatment (239). PBP3 is non-essential and has homology to the PBPs in rod shaped bacteria that work in the PG synthetic complex with RodA, EzrA, and the MreBCD system to elongate the cell prior to cell division (173, 240-242).

PBP4 is the only low molecular weight PBP, and it has roles in creating highly-crosslinked peptidoglycan (172, 243, 244). Furthermore, this PBP is required for glycopeptide resistance and β-lactam resistance, especially in CA-MRSA strains (245-248). The cellular functions of these PBPs remain incompletely understood.

There are many differences between *S. aureus* strains. Though the acquisition of PBP2a distinguishes MRSA from MSSA, other factors encoded by both MRSA and MSSA strains are required to maintain this resistance (139). In addition, some MRSA strains are heterogeneously-resistant to β-lactams, *i.e*, not all the cells in a culture express the same level of antibiotic resistance, while others have high and homogeneous resistance (249). To make matters more complex, there appear to be genetic differences between the MRSA strains that cause nosocomial infections and the MRSA strains most commonly acquired in the community (250). Furthermore, MW2 and USA300 are both community-acquired (CA) strains, but USA300 has become the predominant cause of MRSA infections in the United States and Canada suggesting that there are differences between CA-MRSA strains that impact their success as a pathogen (251-253).

We wanted to validate our claim that the transposition platform described in Chapter 2 would function in any strain and to compare the response of MRSA strains to β-lactam treatment so that we could identify conserved β-lactam resistance factors. Therefore, we created transposon mutant libraries in the MW2 and USA300 strains. Then, we treated the transposon libraries with three different β-lactams that preferentially bind different PBP proteins and performed Tn-Seq. Treating the library with different β-lactams allows us to learn more about the function of each PBP as well as factors that are required for β-lactam resistance in general. These experiments allowed us to discover a new interaction between two genes, *lyrA,* which is of unknown function, and *lytD*, which encodes a glucosaminidase.

3.2 Tn-Seq on MRSA strains identifies interacting proteins

Creating transposon libraries in MW2 and USA300

The platform for Tn-Seq described in Chapter 2 can in theory be used to create transposon libraries in any strain that can be transduced by Φ11 (132). Therefore, to better study MRSA strains, we decided to create transposon libraries in two community-acquired MRSA strains, MW2 and USA300. We chose these strains because they are two of the most clinically-relevant strains in the United States and because neither contains an endogenous Φ11 prophage, which we would need to delete before constructing a library (254, 255).

We still had to make some modifications to these strains before attempting to create a library. The USA300 strain we used (USA300_TCH1516) has a plasmid encoding both erythromycin and kanamycin/neomycin resistance, pUSA300HOUMR (254). We use an erythromycin resistance marker to select for transposon integration and a kanamycin/neomycin resistance cassette on the transposase plasmid, so we require the parent background to be sensitive to these antibiotics. We removed this plasmid using homologous recombination with a pKFC plasmid (256) encoding an in-frame deletion of the erythromycin-resistance gene, and isolation of colonies that had lost the unstable co-integrated plasmid (Figure 16A). We created a transposon library in this resulting strain as described in Chapter 2 (132). On the other hand, MW2 did not have intrinsic resistance to these antibiotics, but the efficiency of transposition was lower than expected (Figure 16B). We hypothesized that this was due to the fact that MW2 is part of a different clonal complex than the other strains. MW2 is a part of CC1 while USA300 and HG003 belong to CC8 (254, 255, 257). The methylation system (*hsdMS*) of *S. aureus* strains belonging to different clonal complexes recognizes and methylates different sites (258). Therefore, when we make the phage lysate in a RN4220 which belongs to CC8 and use it to transduce the transposon into MW2 which belongs to CC1, MW2 recognizes the DNA as foreign

and subsequently digests the majority of it with restriction enzymes before the transposon can insert into the genome. Knocking out the *hsdR* restriction system in MW2 (Figure 16C) (259-261), restored the expected efficiency of transposition, and we were able to use this strain to make a transposon library. Both of these libraries contained the same six donor constructs described in Chapter 2 (132). In case the *hsdR* knockout affected growth with β-lactam treatment, we created a control library in the WT MW2 strain using only the Blunt transposon construct as efficiency of transposition was too low to use all six transposon constructs. For this work, we concentrated on analyzing only the data from the inactivation constructs.
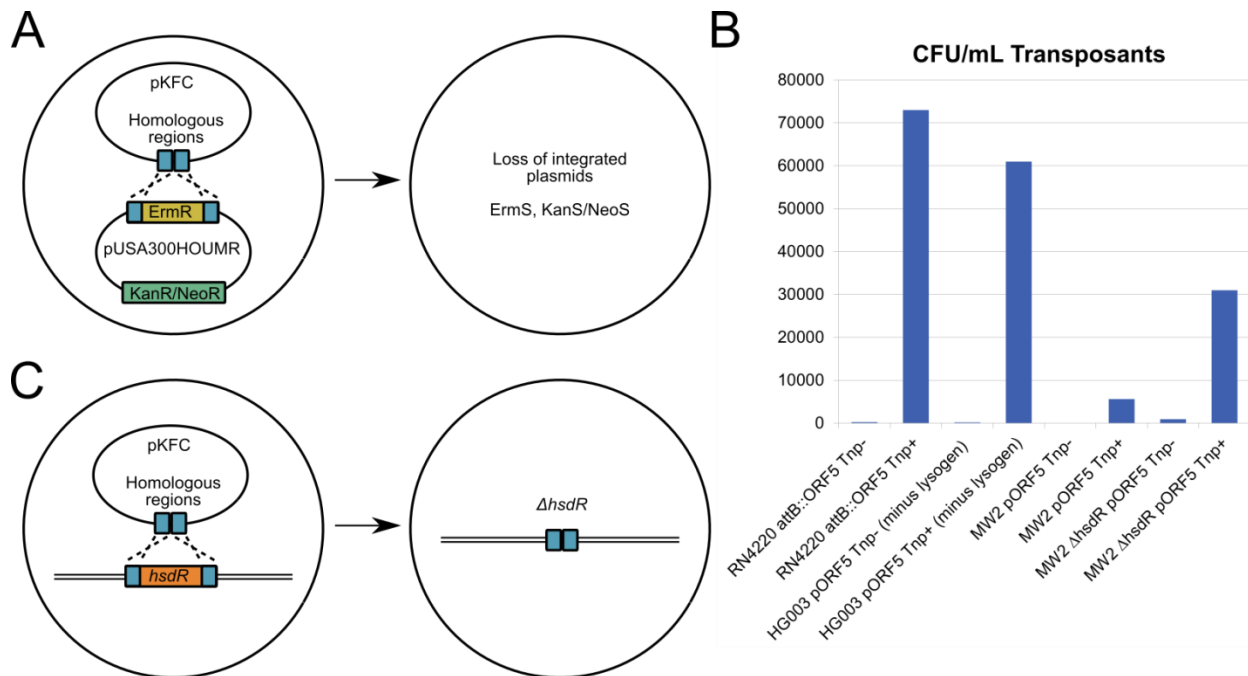
Figure 16. Two MRSA strains were modified for creating transposon libraries. (A) USA300_TCH1516 carries two additional plasmids, one of which encodes resistance to erythromycin and kanamycin/neomycin, pUSA300HOUMR (254). This plasmid was removed via homologous recombination and curing of the co-integrated plasmid. **(B)** The efficiency of transposition for MW2 was lower than expected. Bars 1-4 show transposition effieciency for RN4220 and HG003, while bars 5 and 6 show approximately a 10 fold decrease in transposition efficiency for MW2. **Continued page 61.**

Figure 16 continued. This was calculated by comparing the number of erythromycin-resistant colonies produced from the same amount of transposon-carrying Φ11 phage. To correct this low efficiency of transposition, the *hsdR* restriction system was removed via homologous recombination. Upon *hsdR* deletion, efficiency of transposition was within the expected range (bars 7 and 8).

Treatment of MRSA transposon libraries with β-lactam antibiotics

We chose to treat the transposon library with three different β-lactams that have affinities for different PBPs. Mecillinam binds to PBP3, cefoxitin has highest affinity for PBP4 but can also bind PBP2, and oxacillin has a relatively high affinity for PBP1, PBP2, and PBP3 (Figure 17A) (233). We wanted to treat the library with mecillinam and cefoxitin concentrations where only one PBP was inhibited so that we could learn more about the cellular roles of these non-essential PBPs. We decided to treat the library with low (0.1μg/ml), medium (1ug/ml), and high (10ug/ml) concentrations of oxacillin. The data from the transposon libraries sequenced with low concentrations is useful for identifying factors essential for β-lactam resistance because cells with transposon insertions in genes essential for oxacillin resistance will die and drop out of the library, resulting in fewer reads mapping to these genes compared to the untreated control. Moreover, we will likely not identify many mutations conferring high-level resistance to oxacillin when treating the transposon library at low concentrations of oxacillin because there is not enough selective pressure. However, at the higher oxacillin concentrations, there will be more selective pressure, and there will be few mutants that can survive the antibiotic treatment. In this case, these high fitness mutants will grow better the other mutants in the library, and more reads should map to these genes compared to the untreated control. Therefore, at the higher oxacillin concentrations, it will be easier to identify mutations that can increase resistance to the β-lactam antibiotics.

Traditionally, β-lactam concentrations where only one PBP is inhibited are identified by performing competition experiments using bocillin, a fluorescent β-lactam. However, while these experiments work well using membrane preparations, we wanted to identify β-lactam concentrations where the PBP of interest was inhibited in living cells. We attempted to perform the bocillin competition experiments using cells in mid-exponential phase, but results were inconclusive. Therefore, we devised an alternative approach. We identified the appropriate concentration for mecillinam treatment using data from the temperature screen in Chapter 2 (132). We predicted that *pbp3* would be a temperature-sensitive gene, and therefore, inhibition of PBP3 might also confer the same temperature-sensitive phenotype. Therefore, we compared the growth at 43°C of a *pbp3* mutant (deletion mutant in MW2 from Ambrose Cheung, and inactivation mutant in USA300 from the Nebraska library) (130, 247) to the growth of a WT strain at 43°C treated with different concentrations of mecillinam. Then, we identified the concentration of mecillinam that phenocopied the growth defect of the *pbp3* mutant (Figure 17B). For *pbp4,* it is known that inactivating this gene sensitizes cells to β-lactam treatment. Therefore, we grew MW2 and USA300 with sub-MIC concentrations of oxacillin and increasing concentrations of cefoxitin, and compared their growth to *pbp4* mutants (deletion mutant in MW2 from Ambrose Cheung, and inactivation mutant in USA300 from the Nebraska library) (130, 247). We identified the concentration of cefoxitin that best phenocopied the *pbp4* mutants (Figure 17C). The transposon libraries were treated with these concentrations of oxacillin, mecillinam, and oxacillin, and the sites of transposon insertion were sequenced using Tn-Seq as described in Chapter 2.
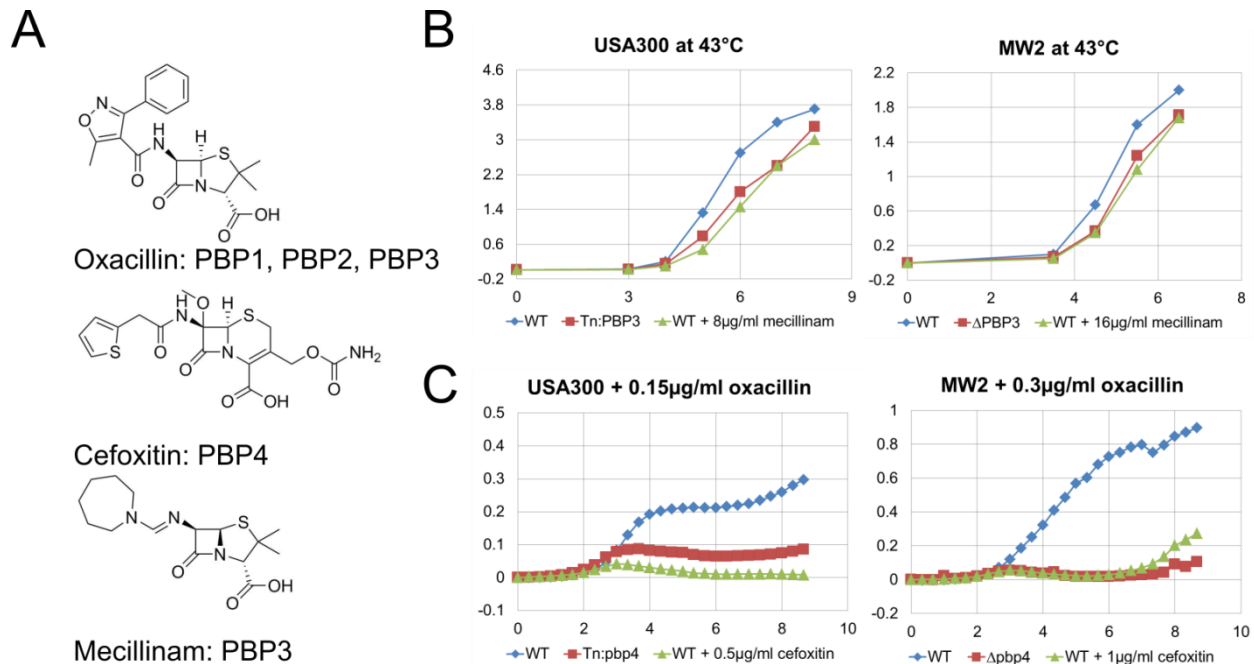
Figure 17. The transposon libraries were treated with three different β-lactams with affinities for different PBPs. (A) Both the MW2 and USA300 libraries were treated with three different β-lactam antibiotics with affinities for different PBPs (233). **(B)** Appropriate concentrations for mecillinam treatment where only PBP3 is inhibited were identified by ascertaining the concentration at which mecillinam treatment phenocopies a *pbp3* deletion strain when both are grown at high temperatures. **(C)** The appropriate concentration of cefoxitin was chosen in a similar way, except that instead of looking for sensitivity to high temperatures, we looked for a concentration of cefoxotin that could potentiate cells to oxacillin in the same way that a *pbp4* deletion did. We have noticed that this strain of USA300 is somewhat more sensitive to oxacillin that MW2. At similar oxacillin concentrations, it grows to a much lower OD600 than MW2. However, we can still identify an oxacillin concentration where the *pbp4* mutation is lethal, and we can identify a cefoxitin concentration that phenocopies the *pbp4* mutant.

Identification of patterns among resistance factors

We analyzed the data as described in Chapter 2 using only the Blunt transposon construct and identified statistically-significant differences in number of reads mapping to each gene using a 10-fold cutoff (ratio in reads/gene > 10 or < 0.1) as a measure of practical significance (Appendix G) (132). We then manually curated the list to remove genes where there were a low number of reads in both the control and the oxacillin-treated condition. This resulted in 41 hits in USA300, 2 hits in WT MW2, and 5 hits in MW2 Δ*hsdR* when the libraries were treated with 0.1µg/ml oxacillin. A greater number of hits in one strain suggests that that strain is more sensitive to oxacillin treatment. We have observed that though these strains have the same growth rate normally, when treating with the same concentration of oxacillin, USA300 grows more slowly and to a lower OD than MW2 (Figure 17C). At 1µg/ml and 10µg/ml oxacillin, there were very few mutants that survived. Because transposon insertions in so many genes were substantially depleted, Appendix G only lists the genes with an increase in number of reads/gene for these two conditions.

Next, we compared the mutations that conferred an increase or decrease in resistance to oxacillin to see if we could use the Δ*hsdR* library, which had a higher efficiency of transposition, as a parent strain for creating transposon libraries. With the exception of *aapA,* an amino acid permease, all the hits in the MW2 libraries were the result in an increase in number of reads that mapped to the gene. To better compare the two MW2 libraries, we needed more genes with a decrease in number of reads after oxacillin treatment. Therefore, we removed the p-value cutoff for depleted genes in the MW2 libraries and gained 11 and 2 hits for the WT and Δ*hsdR* libraries respectively. We compared the data from both MW2 libraries with each other to see if the *hsdR* deletion had any effect by plotting the number of reads that map to each gene in the MW2 WT and Δ*hsdR* libraries for every gene in the untreated condition and the unique set of hits when treated with 0.1µg/ml oxacillin (Figure 18). We observed a linear pattern with a R^2

equal to 0.9815 with oxacillin treatment, suggesting that the *hdsR* deletion strain and the WT strain are very similar, and we can use the higher efficiency Δ*hsdR* strain be used to make transposon libraries.
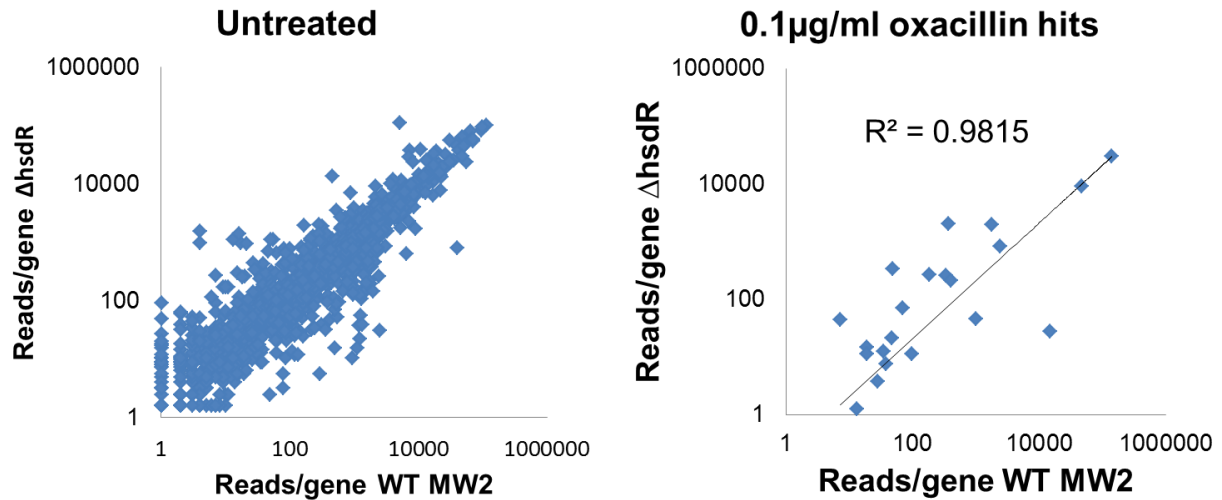
Figure 18. *hsdR* deletion does not have a major effect on the β-lactam resistance factors of MW2. The number of reads that map to each gene for the WT and Δ*hsdR* libraries were normalized to 5 million total reads per library and plotted against each other in the untreated condition (left graph). The same was done for the genes that were hits in either the WT or the Δ*hsdR* library (right graph). In general, the pattern is linear in the untreated condition suggesting that there are no major differences in essential genes when *hsdR* is deleted, and the same trend was observed with the hits when treating with oxacillin (R^2 = 0.9815), suggesting that there are no major differences β-lactam resistance factors as well.

To gain a better understanding of the differences between MW2 and USA300, we identified the top 20 genes that conferred an increase or decrease in fitness when inactivated in the MW2 libraries and compared those to the top 20 genes in the USA300 library (Table 4). Surprisingly, there was only one gene in common between them, *yjbH.* YjbH is a ClpXP adaptor

protein which is known to have roles in disulfide stress and β-lactam resistance (262, 263).

However, there were two other pathways which were represented among both libraries.

Mutations in both purine biosynthesis and menaquinone biosynthesis are known to increase

resistance to β-lactam antibiotics, and in both libraries, we see an increase in reads mapping to

genes in these pathways (Table 4). To identify resistance factors in common between these two

strains, we would have to treat the MW2 library at a slightly higher oxacillin concentration, so

that the number of hits is closer to the number of hits found when we treated with USA300.

These experiments confirm that the genes required for MRSA survival of different stress

responses can vary between different strains. In fact, we were surprised by the extent of the

differences, and these experiments have formed the basis of a comparative genomics approach

that other members of the lab are now using to better understand the differences between these

as well as other hospital-acquired MRSA strains.

Table 4. Hit genes and pathways in MW2 and USA300 when treated with 0.1µg/ml oxacillin			
		Increases Ox resistance in:	
	Gene	**MW2**	**USA300**
Disulfide stress	*yjbH*	X	X
Purine biosynthesis	*pbuX*	X	
	guaB		X
	hpt		X
Menaquinone biosynthesis	*aroA*	X	
	aroC	X	
	menE		X
	menF	X	

We then we identified genes with enriched or depleted reads in common among

antibiotic treatments from each library. Each β-lactam we treated the transposon libraries with

has affinity for different PBPs. Mecillinam has affinity for PBP3, cefoxitin has highest affinity for

PBP4, and oxacillin has similar affinities for PBP1, PBP2, and PBP3. Identifying transposon

mutants that are only sensitive or resistant to cefoxitin or mecillinam treatment allows us to

identify factors that are important for PBP3 or PBP4 function and may help us better understand

the cellular roles of these PBPs. Because oxacillin targets multiple PBPs, we are more likely to

discover genes that generally sensitize to β-lactams with this treatment. Therefore, we looked

for genes with an increase or decrease in transposon reads in any subset of these treatments.

We observed an interesting pattern in the data from the USA300 library. Reads mapping to four

genes, *tarO, mnaA, lyrA,* and *lytD,* were depleted with oxacillin and mecillinam treatment but

were enriched with cefoxitin treatment (Figure 19A). We confirmed that the increase in

transposon insertions was not due to PCR or Bowtie "jackpotting", which results in one TA site

with many reads mapping to it, by plotting the number of reads per TA site in the gene of

interest (Figure 19B). While *lyrA, lytD, and mnaA* all had more reads mapping to the entire gene

with cefoxitin treatment, reads in *tarO* only mapped to the promoter region and the end of the

gene. Though *tarO* is nonessential, it has a significant growth defect where it generally grows to

only an OD of around 0.8, and it has a plating defect where the number of cfu's observed is

always less than the number of cfu's plated. Because of these defects, there are essentially no

knockout mutants of *tarO* present in the library. However, transposon insertions in the promoter

region have the ability to interfere with a gene's native promoter such that the gene is

downregulated, and transposon insertions near the end of a gene could decrease the activity of

the final protein product. Therefore, it is likely that downregulating *tarO* or decreasing the activity

of TarO confers the fitness advantage with cefoxitin treatment that is observed as an increase in

number of reads in these regions. An increase in number of reads mapping to a gene compared

to a control signifies that inactivation of the gene confers either a protective effect or less of a

penalty compared to every other mutant in the library allowing it to grow better than the other

mutants.
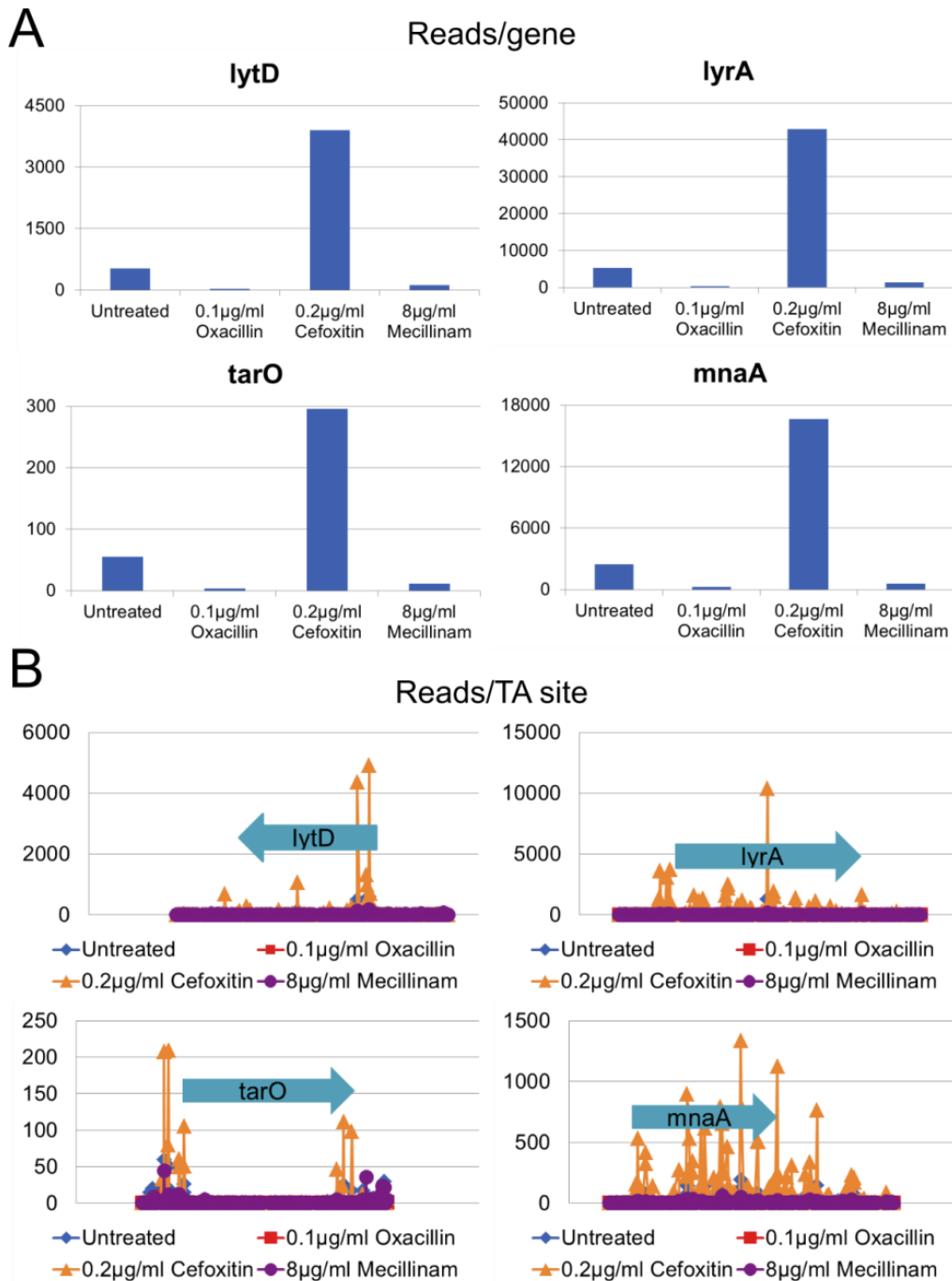
Figure 19. A unique sensitivity when treating with a panel of β-lactams . (A) We looked for genes where the number of reads mapping to the gene was increased or decreased with at least two of the different β-lactam treatments, and we found four genes that when inactivated, produce a similar pattern of resistance and sensitivity: *lyrA, lytD, tarO,* and *mnaA.* **Continued page 69.**

Figure 19 continued. These plots show this pattern after normalization of the data to the same total number of reads. Transposon insertions in these genes appear to confer sensitivity to mecillinam and oxacillin (fewer reads than the control), but resistance to cefoxitin (more reads than the control). **(B)** We confirmed that the increase in number of reads in these genes was not due to "jackpotting", where PCR or mis-mapping results in one TA site with many reads mapping to it. No genes showed evidence of "jackpotting".

TarO catalyzes the first step in wall teichoic acid biosynthesis and it can be inhibited by tunicamycin (264, 265), and *mnaA* is a UDP-GlcNAc epimerase that makes UDP-ManNAc, which is used by another enzyme in the WTA biosynthetic pathway (266). *LyrA,* also known as *spdC*, was initially discovered and named for its ability to confer lysostaphin resistance when inactivated (127). It is a large integral membrane protein with eight transmembrane domains and a large intracellular region at its C-terminus. It most closely resembles CAAX proteases that are involved in membrane anchoring of proteins in eukaryotes (267). However, though *lyrA* has all the conserved residue thought to be the active site residue for the CAAX proteases (127, 268), mutating this residue does not affect its lysostaphin-sensitive phenotype, and LyrA is not thought to have protease activity (127). *LyrA* mutants are also sensitive to tunicamycin (132). *LytD* (also known as *sagB*) was recently shown to be a β-N-glucosaminidase (197). This class of enzymes digests the cell wall between the GlcNAc and MurNAc sugars.

An increased number of reads in a condition compared to the control suggests that inactivation of that gene confers either an increase in fitness compared to the other mutants in the library, or the transposon insertion confers less of a fitness disadvantage in the treated condition than the control. A fitness advantage could be a result of induction of stress responses, removal of a compound's target, or alteration of the flow of metabolites in a pathway. In this case, the concentration of cefoxitin used was non-lethal and cefoxitin's target, PBP4 is

non-essential, which made it surprising that there was enough selective pressure to produce

this increase in number of reads. We hypothesized that some or all of these genes may

physically interact as part of a multi-component machine. In this case, when one component of

the machine is missing or inhibited, the fitness defect would be greater than when the machine

is completely inactivated. To determine whether any of these proteins interacted with each

other, Samir Moussa (a former post doc in the lab) performed co-immunoprecipitation

experiments, crosslinking c-Myc-tagged LyrA to the proteins that it interacts with using DSP

(Figure 20). After purification of the interacting proteins and purification using SDS-Page (this

step also cleaves the crosslink), we observed three bands in both the uncrosslinked and

crosslinked conditions (Figure 20). These proteins were identified by LC-MS/MS, and the

second most abundant protein in the middle band with or without DSP was LytD (Table 5).
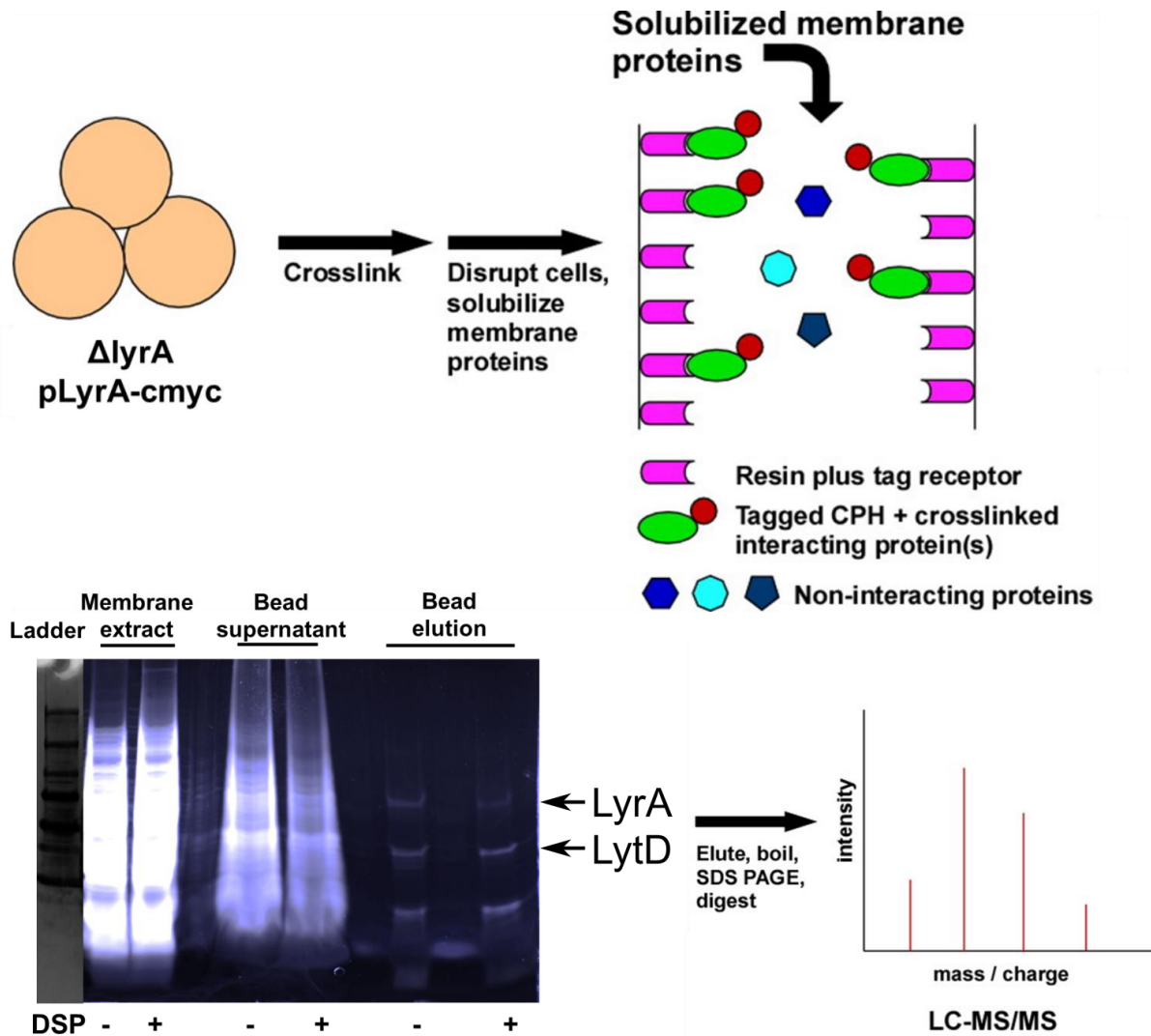
Figure 20. LyrA-interacting proteins were identified using co-immunoprecipitation. A c-Myc tagged copy of LyrA (CPH for CAAX Protease Homologue) was expressed in a Δ*lyrA* strain. A crosslinker, DSP, was added to crosslink LyrA to any other protein it interacts with. Cells were disrupted and membrane proteins were solubilized. These solubilized proteins were purified over a column containing the c-Myc receptor. Specifically-interacting proteins were eluted, purified using SDS-PAGE which also cleaves the crosslinker, trypsin digested, and identified with LC-MS/MS as previously described (269). Addition of the crosslinker did not result in more bands, and the proteins identified by LC-MS/MS were very similar with and without DSP (Table 5).

Table 5. Top 10 LyrA-interacting proteins with highest coverage as identified by co-immunoprecipitation and LC-MS/MS

- DSP			+ DSP		
Top	**Middle**	**Bottom**	**Top**	**Middle**	**Bottom**
SspA	RplB	RplU	SspA	RplB	RplU
SAOUHSC_01659	LytD	RpsM	AtpA	LytD	RplT
GlpD	SAOUHSC_02152	RplT	Tuf	SAOUHSC_01676	RpsE
AtdD	PrsA	RplO	Mqo	RplC	RpsM
Mqo	SAOUHSC_01676	RplV	AtpD	SAOUHSC_02152	RplO
AtpA	SAOUHSC_00634	SAOUHSC_01814	LyrA	SAOUHSC_00634	SAOUHSC_01814
LyrA	RplC	RplS	SAOUHSC_00878	Tuf	RpsF
Fhs	Tuf	RpsE	SAOUHSC_00708	PrsA	RplS
Tuf	SAOUHSC_02699	SAOUHSC_01659	SAOUHSC_00647	SAOUHSC_02366	RpsS
SAOUHSC_00708	Mqo	RpsF	FtsA	SAOUHSC_02699	RpsU

Characterization of *lytD*

If *lyrA* and *lytD* physically interact, their phenotypes to a variety of stressors should be similar. Though there are many unanswered questions about the role of *lyrA,* we do know some of the phenotypes that result from deleting this gene. We know that Δ*lyrA* strains are resistant to lysostaphin (127), sensitive to tunicamycin (95), and that when it is inactivated, sortase-anchored proteins are not properly attached to the peptidoglycan (209). Using a *lytD* transposon mutant from the Nebraska library, we confirmed that strains without functional *lytD* are also resistant to lysostaphin and sensitive to tunicamycin (Figure 21). Furthermore, overexpression of LytD in the Δ*lyrA* background rescues the Δ*lyrA* strain from tunicamycin treatment, and overexpression of LytD in the WT background sensitizes to lysostaphin treatment. This suggests that LyrA may be positivity regulating the activity of LytD, and the phenotype of Δ*lyrA* could be due to misregulation of LytD.
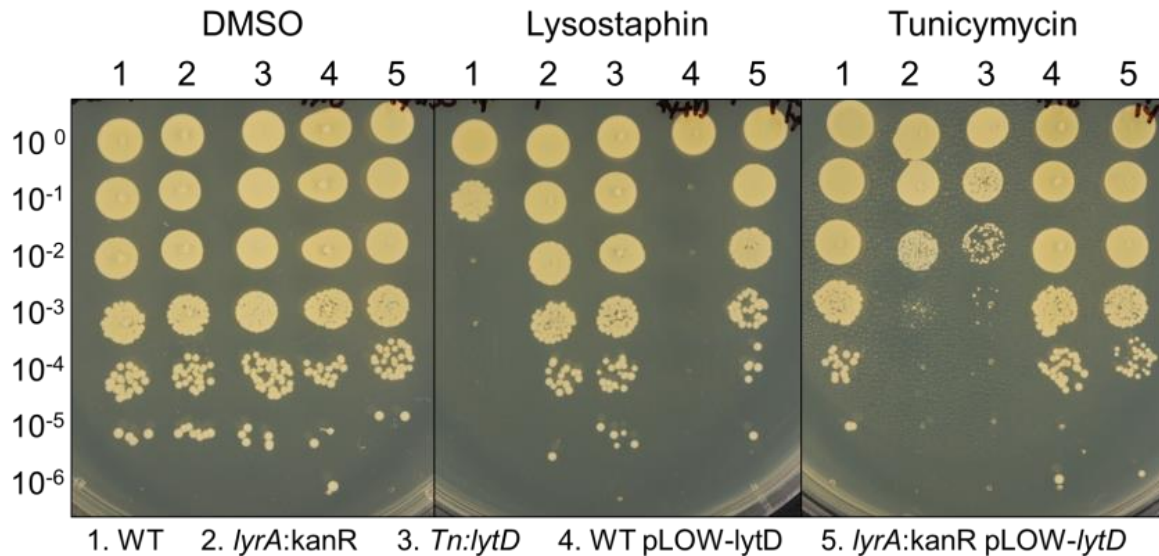
Figure 21. Inactivation of *lytD* phenocopies *lyrA* deletion. Plate dilution spotting assays

were used to test sensitivity of *tn::lytD* to lysostaphin and tunicamycin. Both Δ*lyrA* and the *lytD*

mutant were sensitive to tunicamycin and resistant to lysostaphin. Overexpression of LytD in

WT increases sensitivity to lysostaphin, and overexpression of LytD when *lyrA* is deleted

rescues the sensitivity of *lyrA* to tunicamycin.

Δ*lyrA* strains are reported to have more and thicker cross walls when viewed using TEM

(transmission electron microscopy) (209). We also see an increased number of cross walls in

the *lytD* mutant as well as a WT strain treated with cefoxitin at the same concentration with

which we treated the transposon library, though this could result from a slowing of cell division

when PBP4 is inhibited (Figure 22). Moreover, in the tn::*lytD* strain, we observed many pairs of

dividing cells that are almost entirely separated, but still slightly attached at one end (Figure 22

red arrows). This suggests that *lytD* may have an important role in the final stages of cell

separation. This same phenotype was also observed when WT cells were treated with cefoxitin,

suggesting that inhibition of PBP4 can also produce the same phenotype.
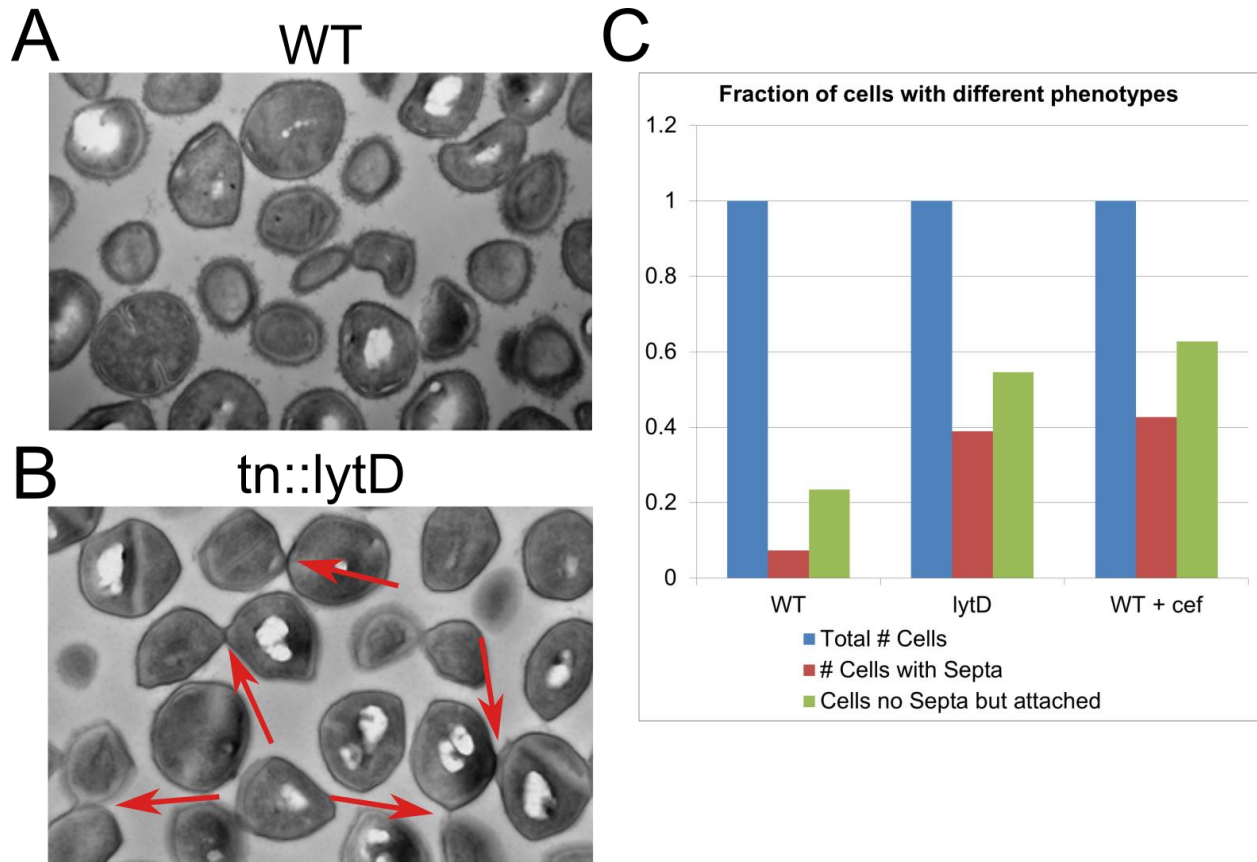
Figure 22. Inactivation of *lytD* results in cells with an increased number of septa. WT cells with and without cefoxitin treatment and *tn::lytD* cells were examined using TEM. Representative photos of WT **(A)** and tn::*lytD* **(B)** cells at 15,000x magnification are shown. We noticed that many of the tn::*lytD* cells had cell separation defects as they were still attached to each other after division had concluded. **(C)** To quantify this difference, we counted the total number of cells, the number of cells with septa, and the number of cells not currently dividing but still attached to each other in TEM images taken at 4,000x magnification (images not shown). Then, we calculated the fraction of the total number of cells with each phenotype. We found that many more *tn::lytD* and cefoxitin-treated WT cells had septa or were still attached suggesting that cell division and separation are slowed.

When *lyrA* is inactivated, there is a significant decrease in the amount of sortase-anchored surface proteins attached to the cell surface (209). Sortase-anchored surface proteins contain an LPXTG cell wall sorting signal which directs it to the cell wall where it is covalently-attached by a sortase (270, 271). We investigated whether *lytD* mutants also had this defect using Staphylococcal surface protein A (SspA) as a marker for proper protein secretion. SspA is an extraordinarily-abundant sortase-anchored surface protein with a YSIRK-G/S signal (271, 272). These signals direct proteins to the cell septum where they are secreted by the Sec system and attached to the cell surface by a sortase. We used antibodies for this protein attached to gold nanobeads purchased from AbCam to observe the localization of SspA in WT as well as *ΔlyrA* and the *tn::lytD* strains (Figure 23). Using TEM, we observed large amounts of SspA on the surface of WT cells, often limited to a few different locations across the surface of the cell (Figure 23, red arrows). In the *ΔlyrA* strain, we sometimes see SspA, but when we do, there is much less than in WT (Figure 23, red arrows), and in the *lytD* mutant, we could not find evidence for any secreted SspA (Figure 23). *lytD* apparently plays a very important role in sortase-anchored surface protein export and/or attachment.
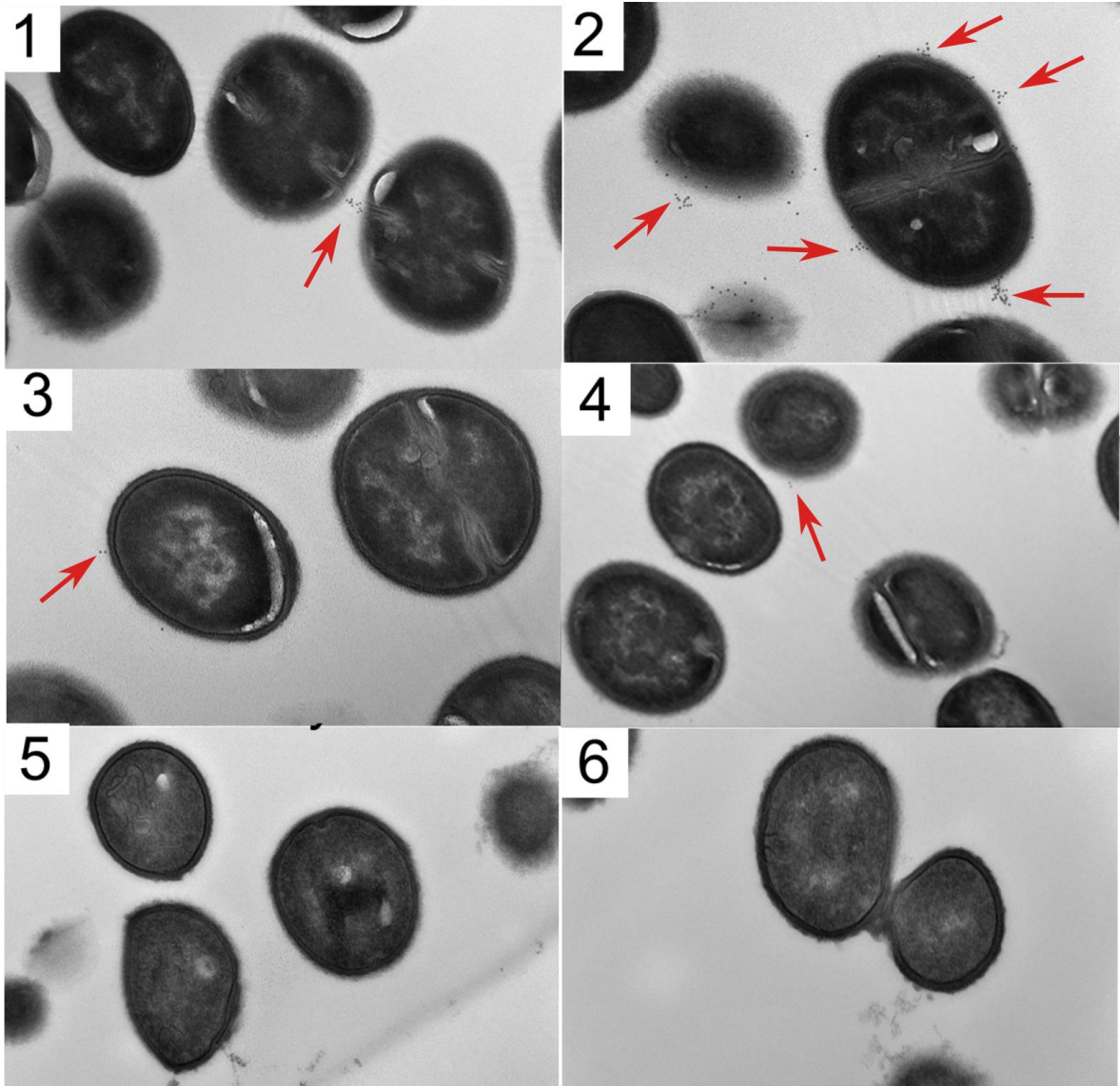
WT: 1-2, ΔlyrA: 3-4, tn::lytD: 5-6

Figure 23. Cells without LyrA or LytD have defects in surface protein export/attachment.

Antibodies to SspA were used in combination with antibodies conjugated to gold nanoparticles to assess how much SspA was attached to the cell surface in WT, Δ*lyrA,* and *tn::lytD* strains using TEM. Representative views of these cells are shown. WT cells had large clumps of SspA, often at the site of cell division (1 and 2), Δ*lyrA* had fewer and smaller clumps of SspA (3 and 4), and *Tn:lyrA* had no detectable SspA (5 and 6). The red arrows point towards clusters of SspA found on the outside of WT and Δ*lyrA* strains.

Sortase-anchored surface proteins have an LPXTG motif and are thought to be attached to the peptide chains of growing peptidoglycan by the sortase enzymes (273, 274). In the absence of *lyrA,* there is a slight increase in resistance to vancomycin which binds to the D-Ala-D-Ala portion of Lipid II, the peptidoglycan precursor (209). Without sortase-mediated protein attachment to the cell wall, there may be more Lipid II available for vancomycin to bind, resulting in more vancomycin required to inhibit peptidoglycan biosynthesis. This was also true for the *lytD* mutant (Figure 24), providing evidence that *lytD* inactivation may be inhibiting protein export through the same mechanism as *lyrA* deletion
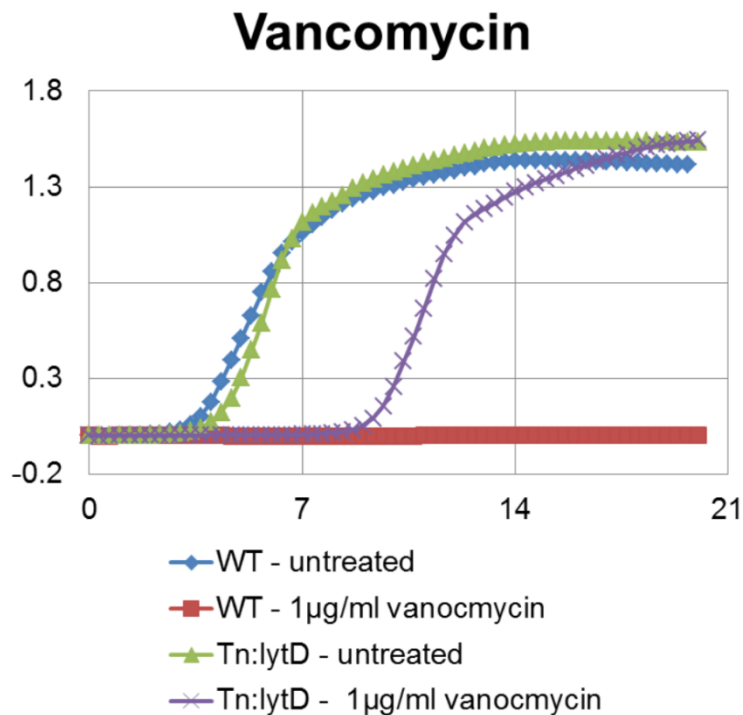
Figure 24. tn::*lytD* mutants are more resistant to vancomycin than WT. *lyrA* mutations are known to confer a small increase in resistance to vancomycin possibly by increasing the amount of extracellular lipid II available for binding to vancomycin (209). We grew WT (TM283) and *Tn::lytD* strains in 96 well plates overnight with vancomycin. We found that *Tn::lytD* is more resistant to vancomycin than WT.

3.3 Discussion

Tn-Seq in MRSA strains

It is very important to perform experiments in the appropriate strain. In our case, as β-lactam resistance is the key characteristic of the epidemic MRSA strains, it makes sense to perform Tn-Seq on MRSA strains. Fortunately, the transposon sequencing platform we have developed allows us to do this while only making minor modifications to the WT strains (254, 255), and these modifications do not seem to affect the resistance factors identified. Another advantage of performing these screens in MRSA strains is that in general, MRSA strains are better able to survive stress, meaning that fewer genes are essential in the control condition. A previous member of the Walker lab, John Santa Maria, performed a Tn-Seq experiment, where he identified the genes which became essential in the absence of wall teichoic acids, by treating a transposon library with tunicamycin and performing Tn-Seq (95). This was how we knew that *lyrA* inactivation conferred a tunicamycin-sensitive phenotype. However, in the strain that screen was performed in, HG003, *lytD* inactivation seems confer a significant fitness defect because few reads map to this gene. If few reads map to a gene in the control, it is very difficult to identify significant depletions in read number in the experimental condition. Because we performed these experiments in MRSA strains, which are better at surviving cell envelope stress, we were able to identify *lytD* inactivation as sensitizing to oxacillin and mecillinam.

Interacting genes have the same resistance and sensitization pattern

We identified four genes, *lyrA, lytD, tarO,* and *mnaA,* that had an increase in number of reads with cefoxitin treatment, but a decrease in number of reads with mecillinam and oxacillin treatment. There are many reasons why genes could have the same resistance pattern to these β-lactams. They could be in the same pathway or in parallel pathways, but in this case, it appears that at least two of these genes have a direct physical interaction. There are many

hypothetical genes in *S. aureus,* and it is likely that many of them have important roles in virulence and antibiotic resistance. This work shows that it is possible to predict interacting genes based on the similarity of the resistance and sensitization pattern across a panel of antibiotics. It is possible that this approach could be applied more generally across any unstudied gene in the genome, though it is likely that we would need more antibiotics with more varied targets in order to identify relevant fitness patterns for other genes that do not have roles in PG synthesis or cell division.

Model for LyrA and LytD relationship

We believe that LyrA is likely regulating the activity of *lytD,* though we do not know the mechanism by which it regulates its activity. One of the major phenotypes of both of these mutants is that they are defective in the export of sortase-anchored surface proteins (197, 209). We can hypothesize a model for this interaction and how it is involved in sortase-anchored surface protein export based on the data described above (Figure 25). This model will help us design future experiments. LyrA is known to localize to the septum (127, 209), and so LytD may localize there as well. *TarO* is required for proper localization of PBP4 to the cell septum, and teichoic acids control the localization of the autolysin, Atl (188, 275). It is possible that TarO or teichoic acids could control the localization of LyrA and LytD to the septum as well. There, LyrA activates LytD to cut the glycan strands. LytD is known to control the length of the glycan strands that are used to create peptidoglycan (197). Glycan strands of a certain maximum length may be necessary for secreted proteins to be exported to the surface of the cell and/or attached to the outside of the cell wall by the sortase enzymes. A major cell separation enzyme, LytN, has the YSIRK-G/S motif, meaning that it is secreted at the cross wall (276). If its export is inhibited when *lytD* or *lyrA* is inactivated, it may explain the defect in cell separation we observed via TEM.
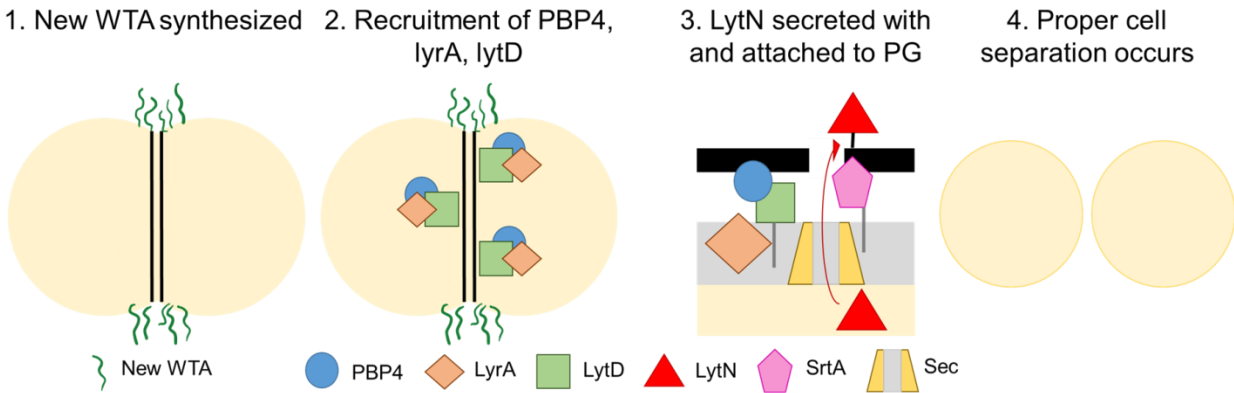
Figure 25. Hypothetical model for LyrA and LytD roles in cell separation. New wall teichoic acids are synthesized at the septum (1). *TarO* and teichoic acids are required for proper localization of PBP4 and Atl to the cell septum, and they may recruit other PG synthesis and cell division factors as well, including LyrA and LytD (2). The secretory machinery exports YSIRK-G/S proteins at the cell septum. LyrA may activate LytD to cut glycan strands. Perhaps long glycan strands inhibit surface protein secretion and/or attachment to the cell wall. LytN is an example YSIRK-G/S-containing protein, which is essential for proper cell separation (3). The proper secretion of LytN allows for successful cell separation (4).

Another major outstanding question is how inhibition of PBP4 in combination with *lyrA* or *lytD* confers an increase in fitness compared to inhibition of PBP4 alone. These results suggest that when PBP4 is inactivated by cefoxitin, the presence of active LyrA and LytD confers a fitness defect. PBP4 is required for highly crosslinked peptidoglycan, and when inactivated, the cell wall is weakened (Figure 26A and B) (231). In the absence of LytD, the glycan strands are longer (Figure 26C) (197). Perhaps longer glycan strands inhibit protein export, but perhaps they can restore some of the cell wall's strength when PBP4 is inactivated (Figure 26C and D). This could explain why inactivation of genes required for proper LytD activity can result in an increase in number of reads (increased fitness) in the presence of cefoxitin. This model (Figures 25 and 26) has allowed us to develop hypotheses for further experimentation.

80
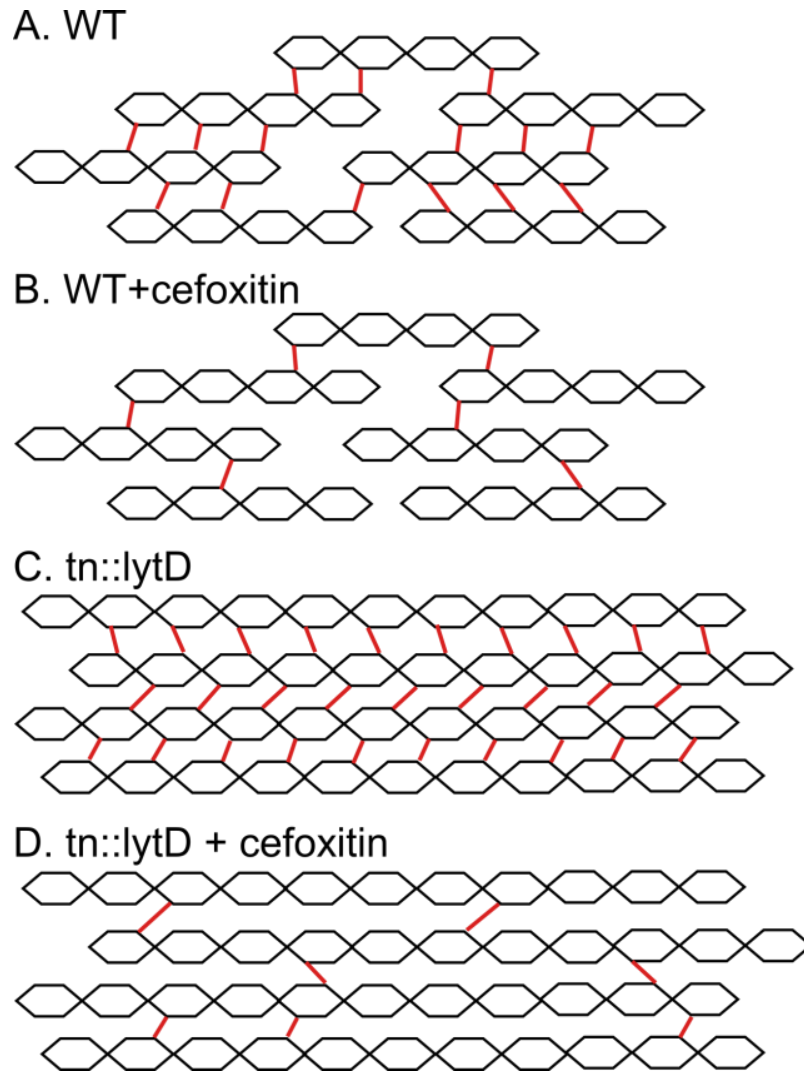
Figure 26. Cefoxitin and LytD affect the structure of the cell wall. A model of peptidoglycan in different strains/conditions is shown. Glycan strands are shown as black hexagons while crosslinking is shown as red lines between glycan strands. **(A)** In a WT cell, LytD controls glycan strand length and PBP4 creates highly-crosslinked peptidoglycan. **(B)** When inactivated by cefoxitin, PBP4 can not create highly-crosslinked peptidoglycan, weakening the cell wall (231). **(C)** LytD controls glycan strand length (197), and when inactivated by a transposon, the glycan strands will be longer. **(D)** Perhaps these longer glycan strands can make up for the decreased crosslinking resulting from cefoxitin treatment, restoring the strength of the cell wall.

Future Directions

Currently, other members of the lab (Katie Coe and Kaitline Schaefer) are studying LytD's enzymatic activity in the presence and absence of LyrA, to see whether LyrA has an effect on LytD activity. It is known at PBP4, TarO, and LyrA localize to the cell septum, but in order to confirm the model above, we need to show that LytD does as well, which can be done using fluorescence microscopy. We also need more evidence to prove that these proteins interact. Furthermore, we have done co-immunoprecipitation using LyrA, but perhaps co-immunoprecipitations using LytD would identify more proteins that interact with this complex.

Conclusion

In conclusion, we were able to show that the method described in Chapter 2 for making and sequencing transposon libraries (132) can be applied to other strains of *S. aureus,* including multiple MRSA strains. We showed that knocking out a restriction system from MW2 increases the transposition efficiency when transducing from a different clonal complex, but does not affect oxacillin resistance factors. Therefore, this strain is a suitable background for transposon library generation. Using cefoxitin, a PBP4-selective β-lactam, as a probe, we identified a possible interaction between LytD and LyrA, and crosslinking experiments confirmed that these proteins physically interact. Both proteins appear to be important for sortase-anchored surface protein export. More work is needed to understand this interaction, but the work so far has allowed us to build a model from which we can design experiments. Finally, we have outlined the beginning of a method for predicting the function of hypothetical genes. By sequencing transposon libraries treated with a wider variety of antibiotics and computationally automating the way similar resistance and sensitization patterns are predicted, we may be able to predict functions, pathways, and even physical interactions for any hypothetical gene.

Chapter 4. Tn-Seq as a tool for studying antibiotics and antibiotic-resistance

We anticipate that contents of this chapter will be published in a journal prior to the publication of this dissertation.

This work was done in collaboration with Mithila Rajagopal and Melissa Martin.

4.1 Introduction

Antibiotic-resistance

Antibiotic-resistant bacterial infections have recently become a global public health crisis, but antibiotic resistance was observed in the clinic only a few years after the first antibiotic was introduced (4). This is due to the fact that in the early years of antibiotic discovery, pharmaceutical companies developed many new antibiotics. The golden age of antibiotic discovery lasted from about 1950 to 1960, and in fact, about half of the antibiotics in use today were discovered during this time (22). Since then, however, development of new antimicrobials has slowed, while bacteria have developed resistance to every clinically available antibiotic (277).

Different antibiotics target different cellular processes by interfering with the function of an essential gene or cellular structure. Some of the cellular processes that are commonly inhibited by antibiotics include cell wall biosynthesis, protein synthesis, DNA/RNA synthesis, and the cell membrane. Inhibition of these processes normally causes cell death, but bacteria have been able to evolve and/or acquire new strategies to increase their resistance to antibiotics. Furthermore, bacteria often have innate mechanisms for up- or down-regulating genes that help defend against antibiotic treatment.

Tn-seq for studying intrinsic resistance factors

S. aureus is a gram-positive pathogen whose ability to evade both antibiotics and the human immune system contributes to the intractability of its infections. Many factors, both internal and acquired, contribute to its ability to survive stressful conditions. *S. aureus* can become resistant to antibiotics by acquiring resistance factors from other species via horizontal gene transfer. For example, VRSA strains have acquired the VanA gene cluster from *Enterococcus faecalis,* which provides vancomycin-resistance to *S. aureus.* In contranst, there can be internal factors, called intrinsic resistance factors, that play a role in the normal physiology of the bacterium as well as contributing to antibiotic resistance. These factors are usually common to both MRSA and MSSA strains.

Individual studies have highlighted the importance of specific intrinsic factors to certain antibiotics. For instance, MprF, a protein that catalyzes the synthesis and subsequent flipping of lysl-phosphatidylglycerol, is a well-studied intrinsic factor, that contributes to resistance to cationic antimicrobial peptides, daptomycin, and some aminoglycosides (206, 278, 279). Also, in MRSA strains, effective pharmacological inhibition of TarO, the first step in wall teichoic acid biosynthesis, with tunicamycin restores full sensitivity to β-lactams even though the β-lactam resistance gene, *pbp2a,* is present (189, 265). These examples show that there are opportunities to mitigate antibiotic resistance by targeting intrinsic resistance factors. Therefore, it is important to identify the full repertoire of intrinsic factors that affect susceptibility to different classes of antibiotics.

Many intrinsic resistance factors in *S. aureus* have been identified by screening mutant libraries for increased susceptibility to a specific antibiotic (51, 280), but there has yet to be a systematic analysis of genes that can confer resistance to multiple classes of antibiotics in *S. aureus*. A global view of all intrinsic resistance factors in *S. aureus* and their effects on antibiotics of various classes will enable further understanding of the bacterial stress response.

A better understanding of these factors will allow for the nomination of candidates for potentiator development, which is particularly important in this "post-antibiotic era".

Recent advances in next-generation sequencing (NGS) have made possible systems-level approaches to identify genes that confer protection to various stresses (89, 132). In this work, we use such an approach to identify intrinsic resistance factors that are important for resistance to six clinically relevant antibiotics having different modes of action. We highlight those that are important across multiple classes of antibiotics and also note factors that confer resistance to certain antibiotics but sensitivity to others. Furthermore, we have characterized two new intrinsic resistance factors for multiple antibiotics which may be important for the cellular response to envelope stress.

Tn-Seq for predicting mechanism of action

We expanded the panel of antibiotics tested to twenty-five antibiotics to use Tn-Seq for predicting mechanism of action of a new antibiotic. Antibiotics target essential cell systems, but not all essential genes make good targets. Even though a gene may be essential, resistance may evolve quickly or it may be difficult for the antibiotic to gain access to the target because of intrinsic efflux genes or membrane permeability issues (281-283). Furthermore, if the antibiotic targets a gene with homology to the human version of that gene, off-target effects in the human body may be a problem. Therefore, before an antibiotic can go in to the clinic, it is helpful to know the mechanism by which it kills the cell. To do this, researchers must figure out the target of the antibiotic. This process can be very lengthy and complex, adding to the cost of developing new antibiotics.

There are a variety of tools that scientists can use to predict mechanism of action. These include, but are not limited to, cytological profiling (284), macromolecular radiolabeling (285), gene expression profiling (286), and proteomics analyses (287). Furthermore, with the advent of next-generation sequencing and the increasing utility of applying machine learning algorithms to

biology, computational approaches for mechanism of action prediction are becoming more common. One method uses an artificial neural network and an antimicrobial peptide's predicted physicochemical properties to predict that peptide's mechanism of action (288). However, this prediction algorithm only works for antimicrobial peptides. Another method, the Antibioticome, utilizes a retrobiosynthetic algorithm along with whole genome analysis to predict a natural product's mechanism of action, but this method would not work with a synthetically-derived antibiotic (289). Finally, another method was recently published that uses a random-forest classifier with a Bayesian learner to predict the mechanism of action of drugs against pathogenic *Saccharomyces cerevisiae* based on their chemical-genetic interactions (290). In addition, this method can predict whether combinations of drugs would synergize or antagonize. This method is powerful, but unless your drug is active in yeast (many antibiotics are not), you cannot use it to predict mechanism of action. Therefore, a computational method that can be used with any class of antibiotic to predict the mechanism of action in bacteria could be very useful for antibiotic development.

Here, we use our high-efficiency phage-based transposon library sequencing platform (132) to develop a two-pronged approach for mechanism of action determination in *S. aureus.* *S. aureus* and its methicillin-resistant relative, MRSA, are dangerous pathogens found as both nosocomial and community-acquired infections (136). We have created ultra-high density transposon libraries in *S. aureus* that have the ability to upregulate as well as inactivate any gene in the genome (See Chapter 2) (132, 133). In order to develop a method for predicting antibiotic mechanism of action in this strain, we have treated a transposon library made in a methicillin-sensitive (MSSA) strain with twenty-five different compounds, including clinically-used antibiotics, agricultural antibiotics, and validated chemical probes, and sequenced the transposon insertion sites using Tn-Seq. The phage-based transposition system utilized here is highly-efficient, allowing us to multiplex different transposon constructs together as described in

Chapter 2. These transposon constructs contain outward-facing promoters of different strengths, and depending on the location and orientation of transposon insertion, they have the ability to constitutively upregulate nearby genes. This is useful to us because target-upregulation is a common mechanism of antibiotic resistance (135). The first method of target-prediction relies on the outward-facing promoters present in our platform to identify upregulated genes. The second method utilizes a machine learning algorithm that compares the mutations conferring resistance and sensitivity to a new antibiotic with the mutations conferring resistance and sensitivity to a panel of 25 known antibiotics. Antibiotics with similar mechanisms have similar resistance factors, so by identifying a known antibiotic with similar resistance factors, we are able to predict the new antibiotic's mechanism of action. This strategy has allowed us to identify and validate both known and novel antibiotic targets and other resistance mechanisms.

4.2 Identification and validation of intrinsic resistance factors

Treatment of the transposon libraries with antibiotics

First, we used a small panel of antibiotics to investigate the intrinsic resistance factors of *S. aureus.* These experiments were conducted in collaboration with Michael Gilmore's lab at Massachusetts General Hospital. We used both our transposon library system (132) as well as data obtained by their lab using a transposon library created using a plasmid-based system (90) and sequenced using the DNA shearing method (291). These experiments were done in different ways (different growth media, growth volumes, and number of generations grown). Therefore, any genes that are identified as hits in both of these methods are more likely to be important for resistance regardless of the conditions in which the cells are grown. Combining these two approaches allows for a more robust dataset that should identify relevant resistance factors to any antibiotic tested. Specific differences between these transposon libraries as well as differences in antibiotic treatment protocols are described in Methods (Appendix A).

The two libraries were treated with sub-minimum inhibitory concentrations (MICs) of antibiotics. These antibiotics - ciprofloxacin, linezolid, gentamycin, oxacillin, vancomycin daptomycin - were chosen for their clinical relevance and selected as representatives of antibiotics that target major pathways: DNA synthesis, protein synthesis, cell wall synthesis, and membrane stability (Figure 27 A-B) (11, 16-19, 292, 293). After treatment, the library DNA was extracted, and Tn-Seq was performed in order to sequence the location of the transposon insertions (Figure 27C).
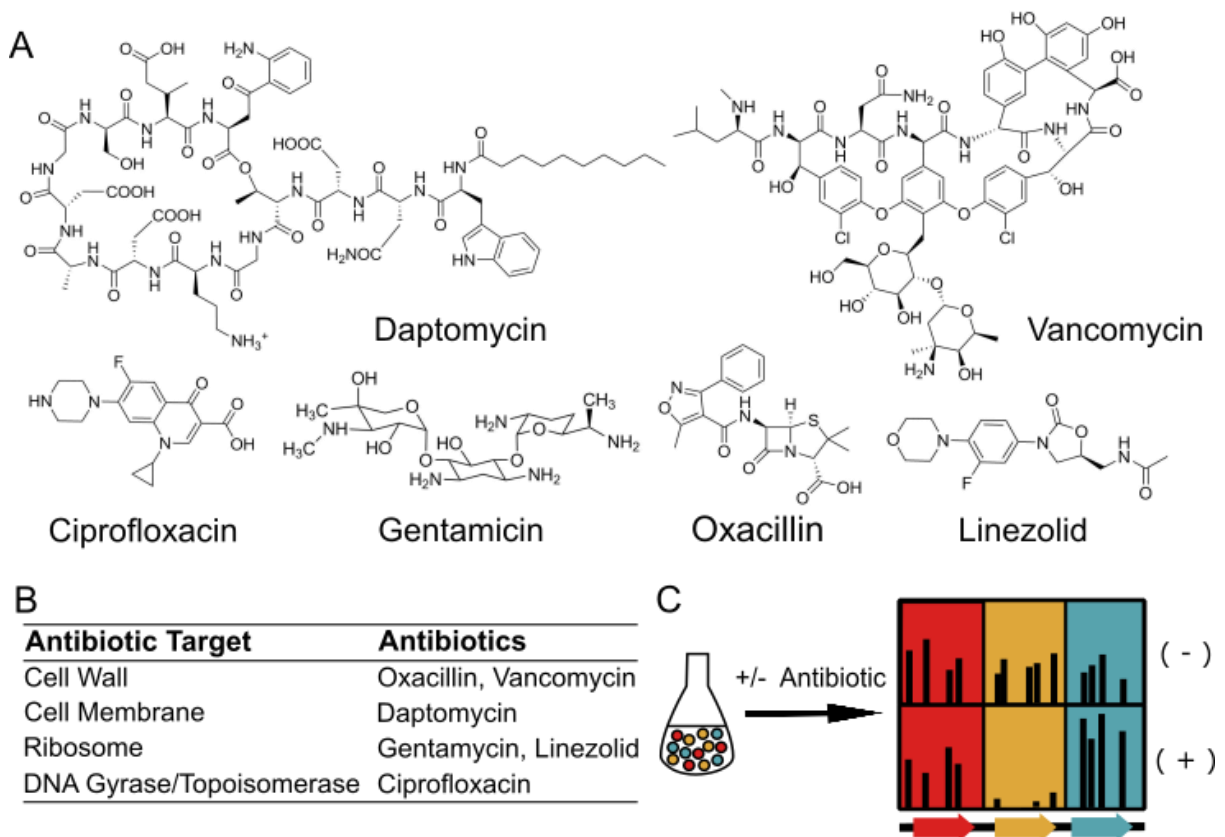
Figure 27. Intrinsic antibiotic resistance factors can be identified using Tn-Seq. The

structures of the six different antibiotics **(A)** and their targets **(B)** used in the Tn-Seq

experiments are shown. **(C)** A pooled transposon insertion library is grown with or without

antibiotic and subjected to Tn-Seq to quantify the number of sequence reads that map to each

insertion location (black lines). Here, the red gene has a similar number of reads in both

conditions. The orange gene has a lower number of reads in the treated sample than in the

untreated control due to a decreased fitness of those transposon mutants in the presence of the

tested antibiotic. Genes of this type are known as intrinsic resistance factors. Finally, the blue

gene has a greater number of reads in the treated sample than in the untreated control signaling

that inactivation of this gene increases bacterial fitness in the presence of the test antibiotic.

Data analysis

Data was analyzed as previously described with one additional step (132). Before comparing the number of reads/gene using the Mann-Whitney U test, the experimental condition (antibiotic treatment) was normalized to the untreated control using simulation-based re-sampling to minimize differences between the two conditions (104, 107). After analyzing all experiments for both libraries separately, the results (ratios and p-values) for all antibiotic treatments were combined. P-values were combined using Fisher's method and the ratio of number of reads in the experiment to the control (hereafter referred to as fold change) was combined by calculating the geometric mean of the fold changes.

Normally, we set one cutoff for statistical significance (p-value) and one cutoff for practical significance (fold change). However, we could not set a single practical significance cutoff for all antibiotics because it would have resulted in different numbers of hits. Different antibiotics seem to exert a different selective pressure on the transposon mutant libraries, even when used at the same fraction of MIC. This results in some datasets where there are large fold changes for a large fraction of the genes, while others had smaller fold changes in fewer genes. Therefore, setting a single cut-off based on the fold-change would result in different numbers of hits for the different antibiotics.

To solve this problem, we set different practical significance cutoffs for each antibiotic, We first identified the subset of genes having p-values $\leq$ 0.05, and then adjusted the fold change cutoff such that a maximum of 20 genes remained. This value was different for each antibiotic, and the cutoffs ranged from a ten-fold increase/decrease in number of reads (10x and 0.1x) for ciprofloxacin to a 55-fold increase/decrease for oxacillin (55x and 0.018x). Then we manually filtered out those genes where all the reads map to only one transposon insertion site as these are likely to be artifacts. It should be noted that this list of top 20 or fewer genes for each antibiotic includes genes with fewer reads in the treated sample compared to the control

as well as genes with more reads (Table 6). The former represent intrinsic resistance factors as disrupting those results in decreased fitness in the presence of an antibiotic, but the latter are also of interest as they provide information on how antibiotic resistance can arise via gene inactivation.

Intrinsic factors that decrease or increase susceptibility to antibiotics

We compared the top hits for each condition and identified 80 unique hits as some genes are found to be hits for more than one antibiotic (Table 6). Twenty-one genes were hits with more than one antibiotic. Fourteen genes were important for resistance to at least two antibiotics, and no gene was a resistance factor for all six antibiotics. These 80 hits include genes involved in almost every major aspect of cell function including cell envelope homeostasis (14 genes), DNA/RNA/protein synthesis (6 genes), protein modification and transport (4 genes), oxidative phosphorylation (11 genes), metabolism/metabolic transporters (11 genes), transcriptional regulators (10 genes), and multicomponent sensory systems (11 genes), Additionally, we have identified thirteen hypothetical genes that are important for resistance.

Whereas the top twenty gene list for most of the six antibiotics included few genes with more reads mapping to them under the treatment condition than the control, half of the genes in the list for gentamycin belonged in this category (Figure 28). All these genes were in the oxidative phosphorylation pathway. It is known that gentamycin and other aminoglycosides rely on the membrane potential to gain entry into cells (294, 295). Disrupting genes in the oxidative phosphorylation pathway therefore limits cellular penetration.

Table 6. Genes with significant changes in number of mapped reads/gene for six antibioticx

	Oxa	Cip	Gen	Lin	Van	Dap
Cell Envelope						
SAOUHSC_00646						
SAOUHSC_00703						
SAOUHSC_00952						
SAOUHSC_00998						
SAOUHSC_01359						
SAOUHSC_01361						
SAOUHSC_01423						
SAOUHSC_01462						
SAOUHSC_01739						
SAOUHSC_01759						
SAOUHSC_02305						
SAOUHSC_02423						
SAOUHSC_02571						
SAOUHSC_03049						
DNA/RNA/Protein Synthesis						
SAOUHSC_01099						
SAOUHSC_01352						
SAOUHSC_01095						
SAOUHSC_00803						
SAOUHSC_01620						
SAOUHSC_01688						
Protein Modification/Transport						
SAOUHSC_00877						
SAOUHSC_01778						
SAOUHSC_01747						
SAOUHSC_01162						
Oxidative Phosphorylation/ETS						
SAOUHSC_01001						
SAOUHSC_01002						
SAOUHSC_00878						
SAOUHSC_00982						
SAOUHSC_01915						
SAOUHSC_01960						
SAOUHSC_01962						
SAOUHSC_01772						
SAOUHSC_01776						
SAOUHSC_02340						
SAOUHSC_02345						
Metabolism/Metabolic Transporters						
SAOUHSC_01708						
SAOUHSC_01040						
SAOUHSC_01043						
SAOUHSC_01611						
SAOUHSC_01852						
SAOUHSC_01803						
SAOUHSC_02552						
SAOUHSC_01430						
SAOUHSC_00536						
SAOUHSC_02360						
SAOUHSC_01013						

	Oxa	Cip	Gen	Lin	Van	Dap
Transcriptional regulators						
SAOUHSC_00023						
SAOUHSC_00467						
SAOUHSC_00503						
SAOUHSC_01228						
SAOUHSC_02362						
SAOUHSC_02298						
SAOUHSC_02299						
SAOUHSC_02300						
SAOUHSC_01979						
SAOUHSC_02303						
Multicomponent Sensory System						
SAOUHSC_00665						
SAOUHSC_00666						
SAOUHSC_00667						
SAOUHSC_00668						
SAOUHSC_02098						
SAOUHSC_02099						
SAOUHSC_02100						
SAOUHSC_02261						
SAOUHSC_02262						
SAOUHSC_02265						
SAOUHSC_01419						
Hypothetical						
SAOUHSC_00468						
SAOUHSC_00774						
SAOUHSC_00788						
SAOUHSC_00965						
SAOUHSC_01025						
SAOUHSC_01050						
SAOUHSC_01568						
SAOUHSC_01569						
SAOUHSC_01645						
SAOUHSC_01724						
SAOUHSC_02149						
SAOUHSC_02228						
SAOUHSC_A02189						

Oxacillin = Oxa   Linezolid = Lin
Ciprofloxacin = Cip   Vancomycin = Van
Gentamicin = Gen   Daptomycin = Dap

Significant increase in reads
No effect
Significant decrease in reads

92

Figure 28 (See page 94). Inactivation of the oxidative phosphorylation pathway confers resistance to gentamicin. (A) Schematic of the oxidative phosphorylation pathway is depicted here. Reads were enriched in the eleven genes named in the figure when treated with gentamicin. A subset of genes in the oxidative phosphorylation pathway were tested to determine if inactivation confers resistance to gentamycin. **(B)** The fold change in reads/gene in the gentamycin-treated sample compared to the control for each of these genes is shown. **(C)** Fitness compared to WT of mutant strains in which the indicated genes were inactivated was assessed using spot dilutions of WT and mutant strains plated on gentamycin (photo on left). Fitness was calculated as the ratio of the highest dilution that allowed growth of WT relative to the mutant (plot on right).

Figure 28 (Continued).
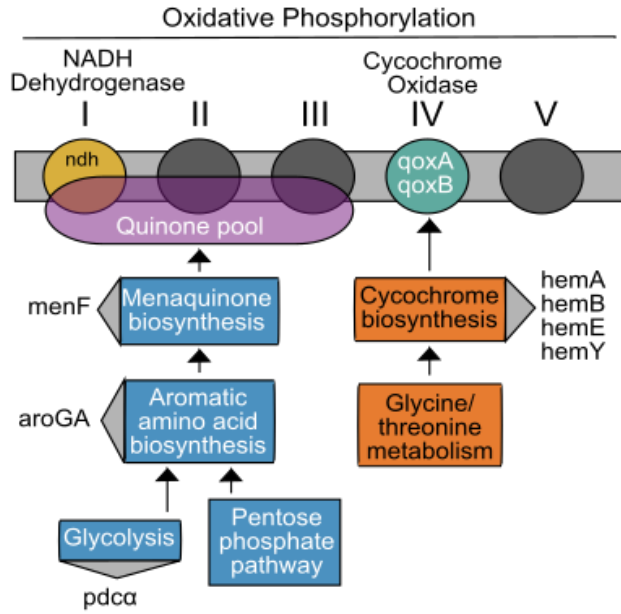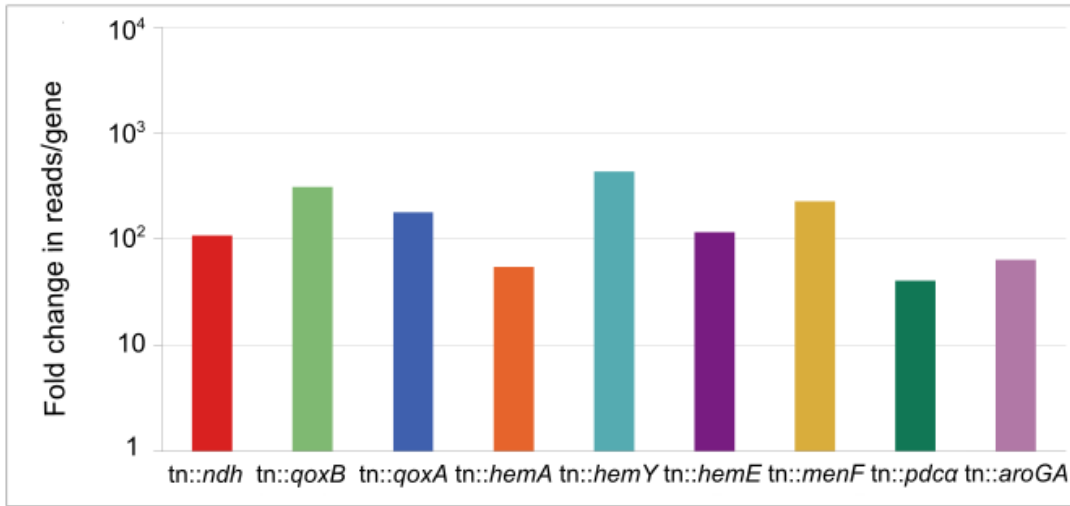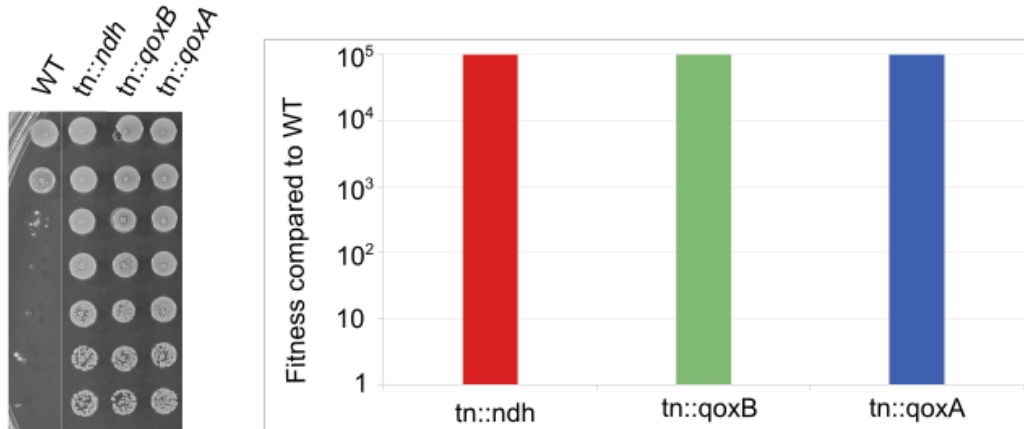
Our fitness analysis also identified many other genes previously known to affect antibiotic resistance. For example, *sigB* was among the hits with oxacillin treatment. Reads mapping to this gene, and the other components involved in the alternative sigma factor pathway, *rsbV* and *rsbW,* were significantly depleted. It has been shown that over-expressing SigB results in cells with thicker cell walls, increased transcript levels of penicillin binding proteins, and elevated MICs to β-lactams (296). Similarly, reads in all three genes of the *vraRST* operon were substantially depleted in the presence of vancomycin. *vraRST* encodes a multi-component-sensing system (MCS) that regulates the cell wall stress stimulon (50, 143, 297). Reads mapping to *pbp4*, a PBP involved in synthesis of highly-crosslinked peptidoglycan and β-lactam resistance (174, 231, 247), were also found to be depleted under oxacillin treatment. Finally, *norA*, which encodes an efflux pump that is known to be involved in ciprofloxacin resistance (7, 298), was identified as an important factor under ciprofloxacin treatment in our analysis. In addition to these and other known intrinsic resistance factors, we have identified 13 hypothetical genes that are important for resistance (Table 6).

Intrinsic resistance factors to multiple classes of antibiotics

Of the twenty-one genes that were hits with more than one antibiotic, eight of these were hits with more than two antibiotics (Table 6). These 8 genes include previously characterized genes - *mprF*, *ndh*, and components of the *graRS/vraFG* multi-component system, as well as two genes of unknown function: *SAOUHSC_01025 and SAOUHSC_01050*. We were encouraged by the fact that we found many previously-identified resistance factors considering the substantial differences between testing an antibiotic against a single mutant and testing it against a huge pooled mutant collection. To provide additional validation of our approach, we tested the fitness of selected mutants (*ndh, fmtA,* and *mprF*) against all six antibiotics using a spot dilution assay (Figure 29). Briefly, this assay works by serially diluting a strain by 10 fold each time, and spotting each dilution on plates containing the antibiotic of interest. A strain more

resistant to the antibiotic than WT will grow at higher dilutions than WT; while WT will grow at higher dilutions than the test strain if the test strain is sensitive to the antibiotic. These three genes were identified as hits with more than one antibiotic treatment. *Ndh* (NADH dehydrogenase) is involved in the electron transport chain (299), *fmtA* is a cell surface protein of uncertain function (300-302), and *mprF* catalyzes formation of lyslyphosphatidylglycerol, a membrane modification that confers protection to cationic antibiotics (51, 190). In general, the agreement between mutant fitness as determined on plate assays is excellent (Figure 29). Notably, the mutants used in the plate assays were made in a different genetic background from the transposon libraries used in the Tn-Seq experiments (*ndh* and *graR* mutants in USA300:FPR3757(130) and *mprF* mutants in Newman while the library was created in HG003 (132)), making the agreement between the fitness values all the more remarkable.

Figure 29 (See page 98). Tn-Seq results were validated by testing mutant fitness in spot dilution assays. (A) Tn-Seq results and validation for selected genes are shown. *ndh* encodes an NADH dehydrogenase involved in oxidative phosphorylation (299); *fmtA* is a cell surface protein of undetermined function (300-302); *graR* is a member of a multicomponent sensing system (MCS) (303); *mprF* and *dltA* are members of this MCS's regulon (304). Top panel: Bar graph depicting fold change relative to the untreated control for each of the six antibiotics. *There are very few insertions in *dltA* control so changes in fitness are not detectable. Bottom panel: Bar graph depicting fitness compared to WT of mutant strains in which the indicated genes are inactivated. **(B)** Fitness was assessed by spotting ten-fold dilutions of WT and mutant strains on antibiotic plates and comparing the highest dilutions that resulted in growth. An example of the plate dilution spotting assays is shown. The left plate has serial dilutions of WT, Δ*dltA,* Δ*graR,* and Δ*mprF* strains spotted on TSA with no antibiotic. On the right, the same strains are spotted on 1µg/ml vancomycin. While growth of WT is unchanged at this concentration of vancomycin, the Δ*dltA* and Δ*graR* strains have severe fitness defects. Δ*mprF* has a more modest fitness defect.

Figure 29 (Continued).

98
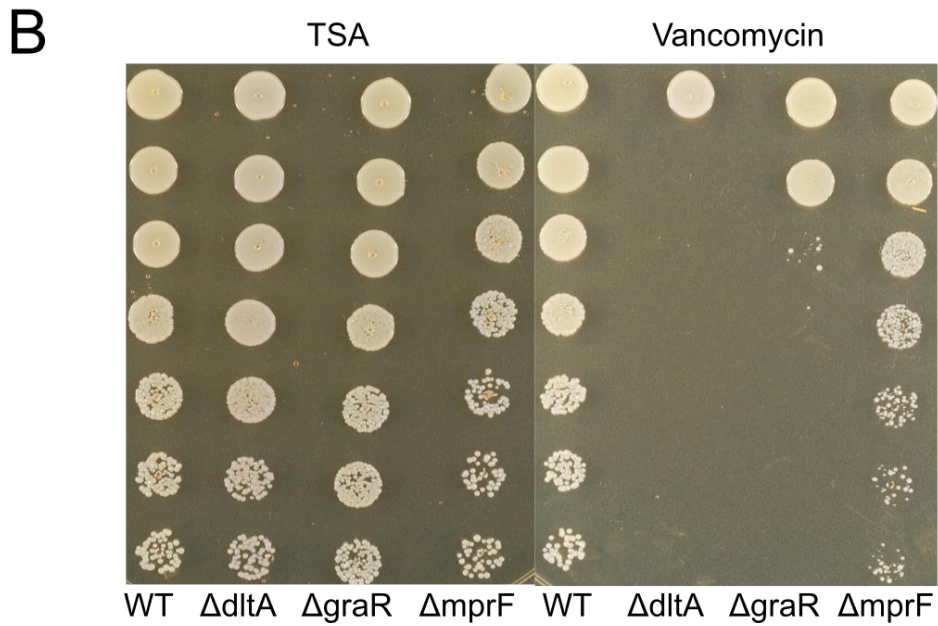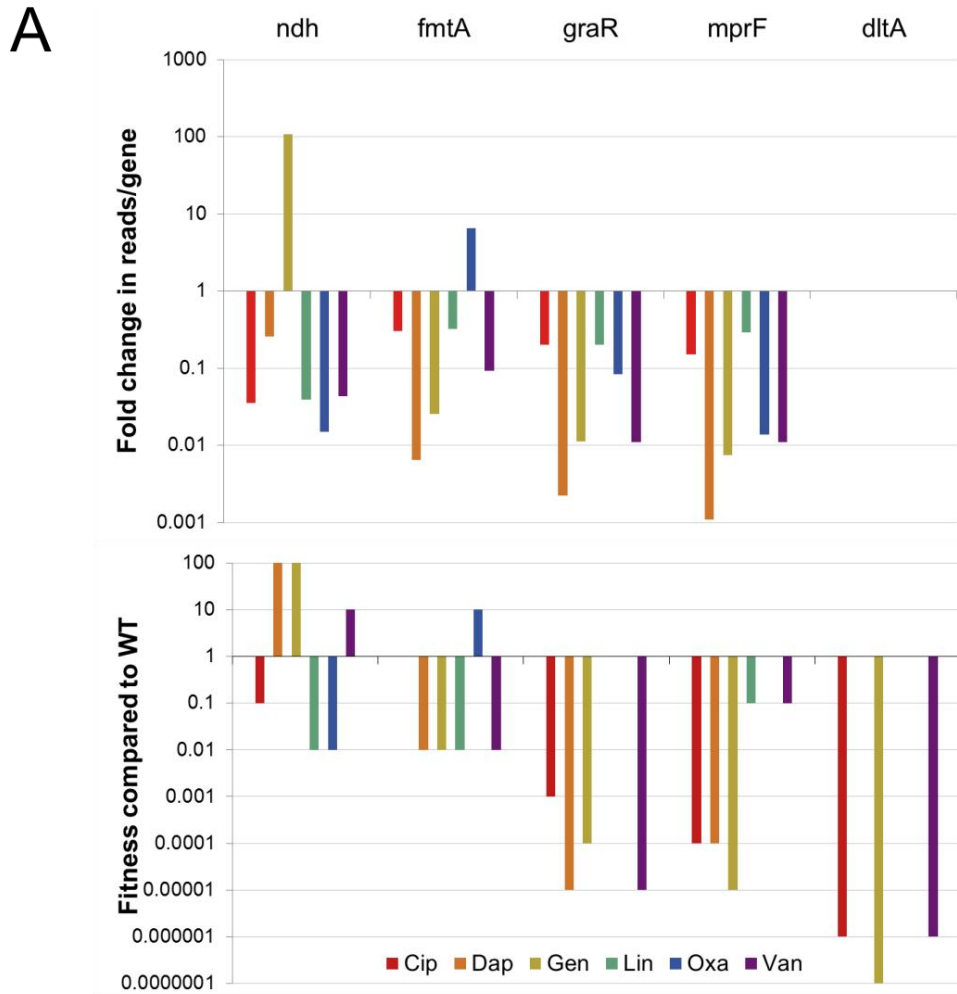
graRS/vraFG is a very important multi-component sensing system

Multi-component sensory/regulatory systems (MCS) allow bacteria to sense and

respond to their environments. These systems typically include a membrane-anchored

extracellular sensory domain fused to an intracellular kinase domain and a separate, cytosolic

response regulator, but they can also include additional elements, such as ABC-transporter-like

and other membrane proteins. A stimulus sensed by the sensory domain results in a change in

phosphorylation of the response regulator, which then modulates the expression of downstream

targets (305). S. aureus contains many multi-component sensing systems, and we identified

multiple components of three of these systems, *agrABCD*, *vraTSR*, and *graRS/vraFG*, as top

hits under treatment with at least one antibiotic (Table 3).

The *agr* locus controls quorum sensing as well as the expression of virulence factors

and autolysins (306-308). We found that reads mapping to *agrA, agrB,* and *agrD* were depleted

under treatment with oxacillin and daptomycin. This suggests that this MCS could be playing a

role in the response to these antibiotics, possibly due to its functions in regulating the autolysin

lytM and the PBPs (144, 309, 310). While *agrC* was also depleted under these treatments, it did

not meet our cut offs.

The *vraTSR* system is known for its crucial role in withstanding vancomycin treatment

(143), and all three components of *vraTSR* were among the top twenty genes identified as

important under vancomycin treatment. This sensing system regulates expression of cell wall

biosynthetic genes and has also been implicated in β-lactam resistance (50, 144, 297). Although

reads mapping to these genes were also depleted under oxacillin treatment, they did not meet

our cut offs for top 20 most important genes.

The single most important MCS across all the six antibiotics tested was found to be

graXRS/vraFG. Four components of this system made our cut offs under gentamycin,

daptomycin and vancomycin treatment . Moreover, compared to wildtype, we found the fitness

99

of a Δ*graR* mutant to be reduced by four to five orders of magnitude when plated on these antibiotics (Figure 29).It was also sensitive to ciprofloxacin, although less so. The *graRS/vraFG* reguon was characterized and found to have links to other global regulatory systems such as *agr* and *walKR* as well as many stress-response and virulence genes (304), but it is most well-known for modulating the charge of the cell envelope through modulating the expression of *mprF* and the d*lt* operon. MprF attaches lysine to phosphatidylglycerol (190, 311) while the *dlt* operon attaches D-alanine to lipo- and wall-teichoic acids (312). These pathways confer protection to cationic antimicrobial peptides, aminoglycosides, and other positively-charged antibiotics (313). Whereas mprF was identified as a top hit under several treatment conditions, transposon insertions in the dlt genes were poorly represented in the control libraries because *dlt* mutants have substantial fitness defects and do not compete well with other strains. The fact that this system as well as a member of its regulon is so prevalent among the top hits was surprising, and it suggests that this system may be more important for antibiotic resistance than previously appreciated.

We next asked whether the sensitization observed when knocking out *graRS/vraFG* is due to misregulation of the expression of *mprF* or *dltA.* We found the fitness of a *dltA* mutant to be greatly reduced compared to wildtype when plated on four of the six tested antibiotics (Figure 29). Three of these, vancomycin, gentamicin, and ciprofloxacin, contain at least one net positive charge. The fitness of the *graR* mutant was also reduced on these antibiotics, but there was not a strong correlation between the number of positive charges and the fitness of the *dltA* mutant across different classes of antibiotics. For aminoglycoside antibiotics it was previously shown that the number of charges correlates with activity against the *dltA* mutant (105). D-alanylation serves different roles in the cell envelope (313-315), and so the fitness of the *dltA* mutant in the presence of different antibiotics may reflect these various roles. Similarly, the data suggest that MprF does not function simply to modulate cell membrane charge (51) as the *mprF* mutant is

most sensitive to daptomycin (Figure 21), which does not contain a net positive charge. Notably, the *graR* mutant was also sensitive to daptomycin, but the *dltA* mutant was not (Figure 29). It seems that depending on the type of stress this MCS senses, it may modulate the expression of many different genes including, but not limited to, *dltA* and *mprF.*

Identification of two new resistance factors important in envelope stress

Of all the hits, thirteen are uncharacterized hypothetical genes. Two of these genes, *SAOUHSC_01025* and *SAOUHSC_01050,* were predicted to be resistance factors for the cell-envelope-targeting antibiotics oxacillin, vancomycin, and daptomycin. These genes are conserved in *S. aureus* and are both predicted to be conserved polytopic membrane proteins. According to the TMHMM Server v. 2.0 which predicts transmembrane helices in proteins, SAOUHSC_01025 is predicted to have six transmembrane helices, a 93 amino acid extracellular domain, and four more transmembrane helices. SAOUHSC_01050 is predicted to have three transmembrane helices and a large 191 amino acid extracellular domain. A BLAST search with SAOUHSC_01025 resulted in no homology to any characterized protein (all were annotated as hypothetical). So, a PSI-BLAST search against NCTC_8325, which is designed to detect distant sequence similarities was performed. This search reveals that SAOUHSC_01025 has low (26%) identity with ComGB. In *Streptococcus mutans* and in *B. subtilis*, ComGB is a protein required for DNA uptake as well as biofilm formation (316), but its role in *S. aureus* has not been studied. PSI-BLAST was also used to identify *S. aureus* genes with homology to SAOUHSC_01050. We found that this protein has high homology (99% identity) with *S. aureus* proteins annotated as chitinases (Figure 30). Chitinases in bacteria are not well studied, but they have been shown to degrade colloidal chitin and peptidoglycan (317). Gene and protein annotations in *S. aureus* are not always accurate, but the similarity of SAOUHSC_01050 to a class of enzymes that has activity on peptidoglycan suggests an important role for this gene in cell envelope homeostasis.

Figure 30. SAOUHSC_01050 has high homology with *S. aureus* chitinases. PSI-BLAST

was used to identify *S. aureus* proteins with homology to SAOUHSC-01050 (highlighted in

yellow). Four of the top ten proteins identified were annotated as chitinases, which

SAOUHSC_01050 has 99% homology with. We aligned three of these chitinases to

SAOUHSC_01050 using Clustal Omega, and we confirmed that these are likely the same

protein because there are differences in only three amino acids. These three amino acids are

shown with colons instead of asterisks.

We were able to mine the Tn-Seq data to try to better understand which pathways *SAOUHSC_01025* and *SAOUHSC_01050* are involved in. Our strategy here was based on the observations that we made in Chapter 3 where we identified at least two physically-interacting genes by identifying genes with the same resistance and sensitization patterns across a panel of β-lactams. In these experiments, we observed that inactivation of genes in the same pathway conferred a similar level of sensitization to each antibiotic. For example, every component of the *graRS/vraFG* MCS were resistance factors to the same subset of antibiotics: gentamicin, daptomycin, and vancomycin. Therefore, we hypothesized that we could nominate pathways for genes of unknown function by identifying known genes with very similar patterns of sensitization to these six antibiotics.

One way to identify similar patterns in an unbiased way is to take advantage of machine learning algorithms (Figure 31A. We optimized a K-nearest neighbors machine learning algorithm using the Sci-kit learn Python library (318) to find genes that have the most similar resistance patterns to our gene of interest. To do this, we calculated the "fitness value" (See Methods) for each gene in each of the antibiotic treatments. In short, this value is based on the change in number of reads mapping to a gene, but it also normalizes for gene length and for insignificant changes when few reads map to the gene in the untreated condition. We validated our algorithm by using the components of *graXRS/vraFG* as test cases, to see if the other components could be identified among the top 5 nearest neighbors (top 5 most genes with similar patterns) when we searched all the non-essential genes in the genome (2549 genes) (Figure 31A). We found that for each of these genes, at least two of the other components of the system were among the five closest genes in the genome, with the exception of *graR* (Table 7). *GraR* does not identify other components of this MCS because it often has a different fitness value than the other member of the MCS (Table 7, Figure 31B). With daptomycin, gentamicin, oxacillin, and vancomycin treatment, the fitness value for *graR* is approximately 10 fold higher

than for the other components of the system. Therefore, it makes sense that this algorithm would not identify other components of this MCS when searching with *graR. GraR* may have a slightly different fitness profile than the other components of this MCS because it is the transcriptional regulator of the system. If *graS, vraF,* or *vraG* are knocked out, the cell envelope stress sensing portion of the MCS, all of which resides in the membrane, may be somewhat defective, but *graR* could still retain some activity. If *graR* is knocked out, there is no way this system can modulate the expression of members of its regulon.
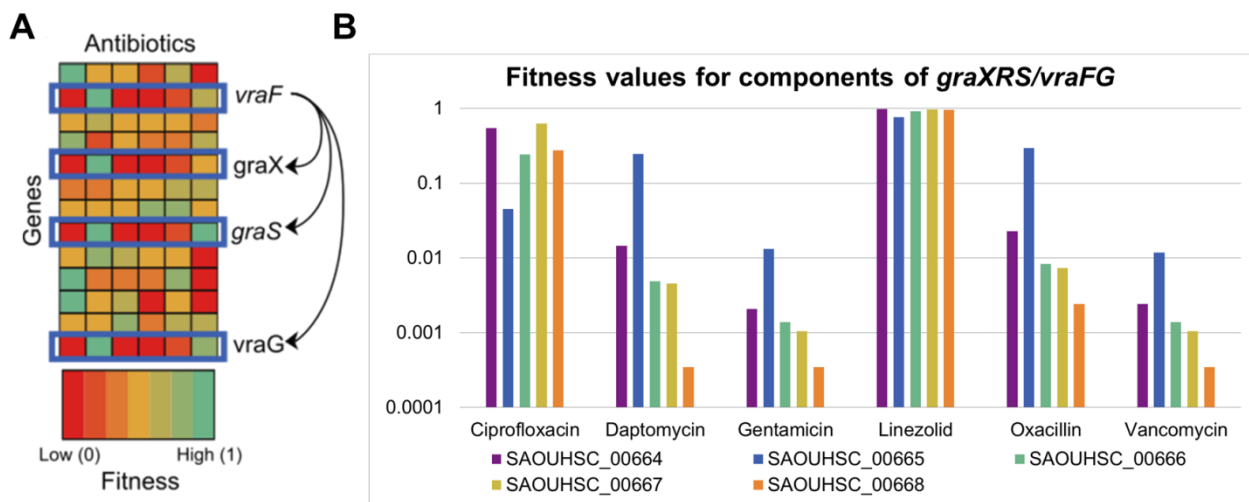
Figure 31. The K-nearest neighbors algorithm can be used to identify genes with similar resistance patterns. (A) Schematic depicting the fitness of a subset of genes upon treatment with different antibiotics. Each column represents an antibiotic and each row represents a gene. Genes having related functions, such as the components of the GraRS/VraFG MCS, have similar fitness profiles across a panel of antibiotics. For any given test gene, we can identify genes with the most similar fitness profiles. These genes likely have functions that are related to the function of the test gene. **(B)** Each component of the *graXRS/vraFG* MCS was put through this analysis to identify the top five genes with the most similar pattern of resistance and sensitivity to these six antibiotics. With the exception of *graR,* every component identified at least two other components of the MCS.

Figure 31 Continued. The fitness values for *graXRS/vraFG* were plotted. With daptomycin, gentamicin, oxacillin, and vancomycin, *graS, vraF, vraG,* and sometimes *graX* had very low fitness values. In contrast the fitness value of *graR* was on average 10 fold higher.

Table 7. Five genes with most similar resistance patterns as identified by K-nearest neighbors

Test Gene	*graX*	*graR*	*graS*	*vraF*	*vraG*	*1025*	*1050*
Top 5 Nearest Neighbors	graS	gdhA	graX	graX	graX	graS	graR
	vraF	1050	vraG	graS	graS	vraG	graS
	vraG	cvfC	1025	vraG	1025	1050	1025
	cydD	2235	1050	cydD	1050	mprF	mprF
	fmtA	murA	mprF	fmtA	mprF	cvfC	cvfC

Next, we applied this machine learning algorithm to our unknown genes of interest *SAOUHSC_01025* and *SAOUHSC_01050.* The five genes most similar to *SAOUHSC_01025* were *graR, graS, mprF, cvfC,* and *SAOUHSC_01050*, while the five most similar to *SAOUHSC_01050* were *graS, mprF, cvfC, vraG,* and *SAOUHSC_01025* (Table 7)*. CvfC* (conserved virulence factor C) was identified as a gene required for virulence in a silk worm model of *S. aureus* infection (319). The function of *cvfC* is not known, but when this non-essential gene is deleted, it confers attenuated virulence in mice and decreased hemolysin production (319, 320). The presence of *graS, mprF,* and *vraG* in the results of this analysis suggests that they have functions related to maintaining membrane integrity and withstanding envelope stress.

Characterization of *SAOUHSC_01025* and *SAOUHSC-01050*

Using spot dilution testing, we tested the fitness of mutants with inactivating transposon insertions in *SAOUHSC_01025* (tn::1025) and *SAOUHSC_01050* (tn::1050) against a panel of twelve antibiotics with different targets (Figure 32). In addition to the six antibiotics originally used to probe the transposon libraries, we tested moenomycin, targocil, bacitracin, fosfomycin, mupirocin and rifampicin (Figure 32). Moenomycin and bacitracin inhibit peptidoglycan synthesis by interacting with extracellular targets (321, 322), whereas fosfomycin inhibits peptidoglycan synthesis by interacting with an intracellular target (323). Targocil inhibits wall-teichoic acid biosynthesis, resulting in depletion of peptidoglycan precursors and, therefore, inhibition of peptidoglycan synthesis (324). Mupirocin inhibits protein translation by targeting an acyl-tRNA synthetase (325), and rifampicin inhibits RNA polymerase (326). With one exception, both mutants showed moderate and similar reductions in fitness when plated on non-cell wall related antibiotics. The exception was that tn::1050 was far more sensitive to gentamycin than tn::1025, an unusual pattern not observed with any other antibiotic. Gentamicin enters the cell through the membrane. As a membrane protein, 1050 may play a role in gentamicin's entrance to the cell, such that when it is inactivated, gentamicin can enter the cell more easily. Against cell wall active antibiotics, both tn::1025 and tn::1050 showed large reductions in fitness, with tn::1025 often found to be more susceptible than tn::1050. Moenomycin provides a striking example of this as the fitness of the tn::1025 mutant decreased by four orders of magnitude while the fitness of the tn::1050 mutants did not change. In general, however, both mutants showed reduced fitness compared to the wild type strain. The decreased fitness of both mutants when plated on cell wall active antibiotics compared with other antibiotics is consistent with an important role in cell envelope integrity (Figure 32).
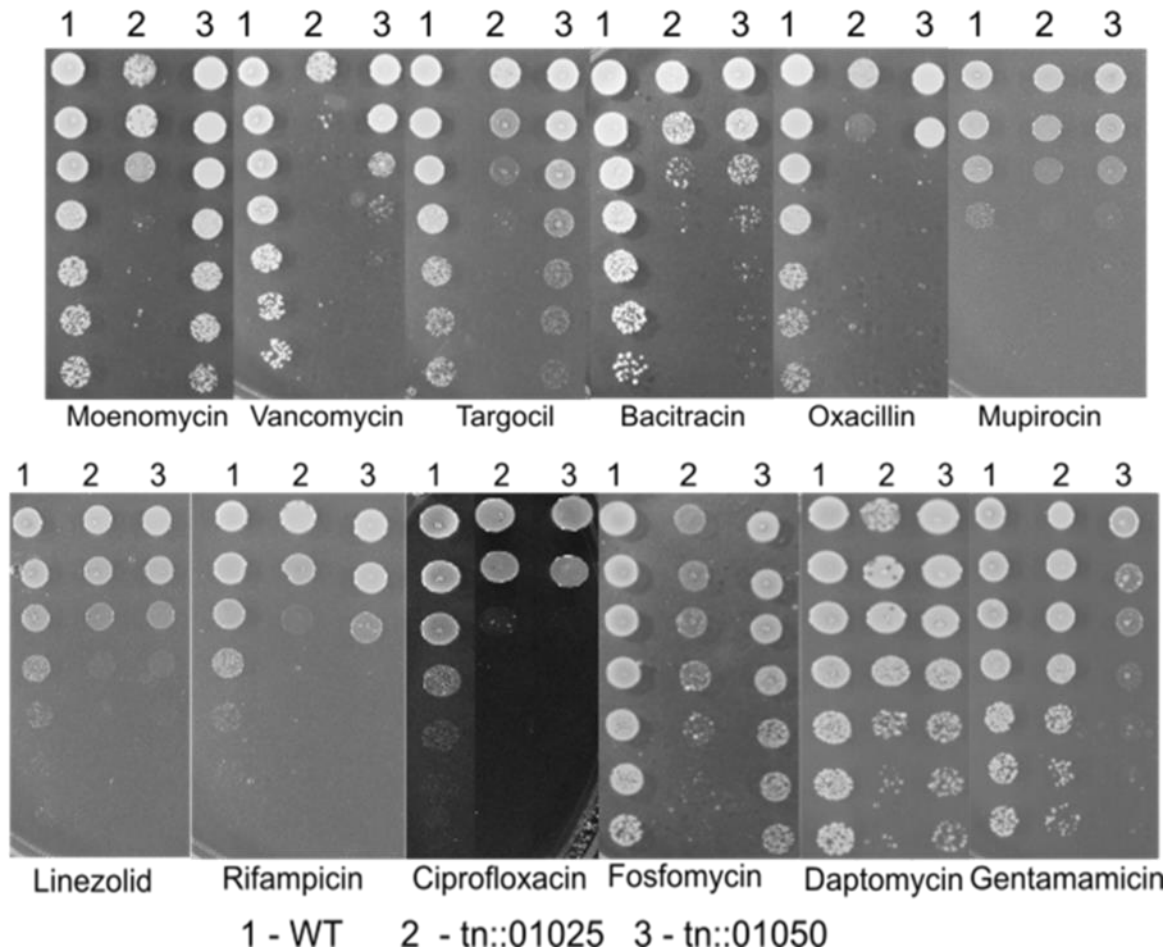
Figure 32. Two genes encoding polytopic membrane proteins are important for resistance to cell envelope targeting antibiotics. The K-nearest neighbors algorithm predicted that *SAOUHSC_01025* and *SAOUHSC_01050* were most similar to one another and also shared three of four other identified neighbors. As these neighbors play an important role in protecting S. aureus from certain classes of antibiotics, we predict that *SAOUHSC_01025* and *SAOUHSC_01050* are important for cell envelope integrity. Spot dilution assays showing fitness compared to WT of inactivation mutants in *SAOUHSC_01025* and *SAOUHSC_01050* upon plating on the indicated antibiotics. **Continued page 108.**

Figure 32 continued. The first five antibiotics (moenomycin, vancomycin, targocil, bacitracin, and oxacillin) target the cell envelope and at least one of the two mutant strains is highly sensitive at an antibiotic concentration that permits growth of WT at all dilutions. With the exception of fosfomycin and daptomycin, the rest of the antibiotics have non-cell envelope targets, and we observe either little effect on these mutants (fosfomycin and daptomycin), or we observe an effect only at concentrations where the WT strain is also inhibited (mupirocin, linezolid, rifampicin, and ciprofloxacin). The last antibiotic is gentamicin, which inhibits protein synthesis but enters the cell using the membrane potential. Transposon insertions in *SAOUHSC_01050* sensitize to this antibiotic.

These mutants did not show a significant increase in sensitivity to lysis by triton X-100 (Figure 33), but they showed a slight reduction in growth rate at 37°C. The doubling times for WT, tn::1025 and tn::1050 were 22.0, 25.4, and 26.5 minutes respectively, suggesting that these mutants have some fitness defects. Though a three or four minute increase in doubling time, may seem unsubstantial, in the context of a transposon library in liquid culture, where all the mutants are competing against each other for nutrients, it is easy for a small fitness defect to become amplified over successive generations.
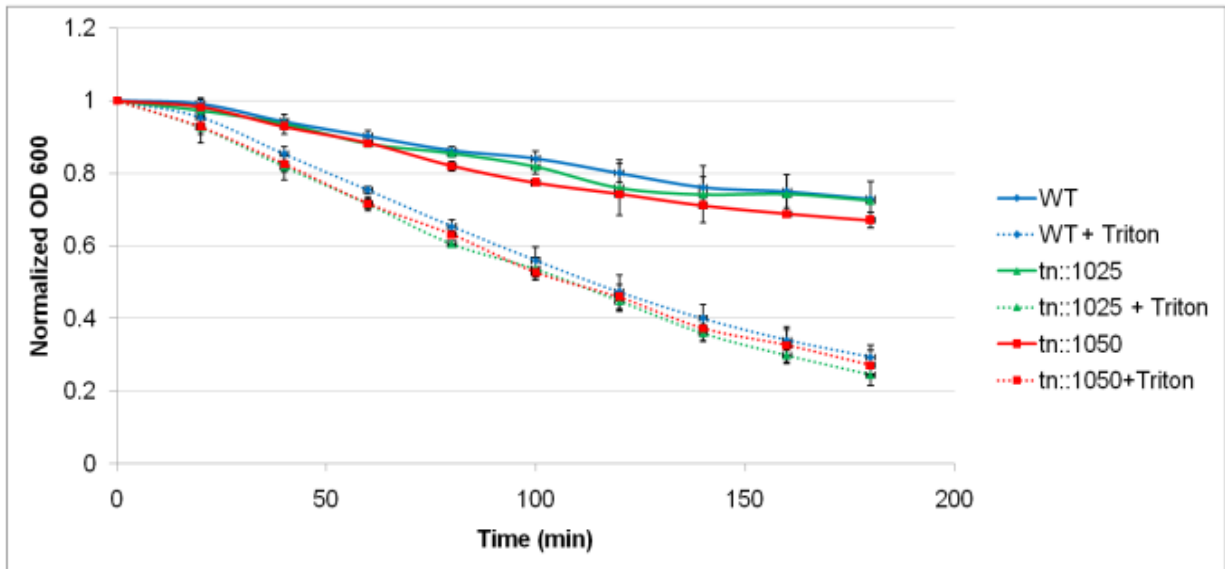
Figure 33. Inactivation of *SAOUHSC_01025* or *SAOUHSC_01050* do not have an effect on TritonX-100 induced lysis. Lysis curves of WT, tn::*1025* and tn::*1050* with or without triton X-100 show that WT and mutants lyse at similar rates. OD600 values were normalized to the initial starting OD of each sample.

Because the antibiotic resistance and sensitization patters of 1025 and 1050 were found to be similar to *cvfC* (319, 320), we tested whether there was any change in hemolysin production and activity by streaking WT, tn::1025 and tn::1050 on Columbia Blood Agar plates. In contrast to the *cvfC* mutant, which shows reduced hemolysin production, these mutants had greater hemolysin production than the WT strain (Figure 34). *S. aureus* encodes four hemolysins, α-, β-, γ-, and δ-hemolysin (34). Expression of the hemolysins is tightly regulated by the two component system SaeRS and the Agr system (327-331). In the presence of heme, α- and β-hemolysin expression is down-regulated to protect the cell from the toxic byproducts of heme degradation (327), while in blood, γ-hemolysin expression is upregulated (332). Because the tn::1025 and tn::1050 mutants increase hemolysin expression, it is possible that they function to negatively regulate their expression or activity. Though inactivating these genes

upregulates hemolysin production and *cvfC* knockouts have less hemolysin production, it still makes sense that we could identify *cvfC* as having a similar antibiotic resistance and sensitivity pattern. Depending on the type of cell stress, different transcriptional regulators can up- or down-regulate the expression of different sets of genes. Therefore, under cell envelope stress due to antibiotic treatment, inactivation may confer a similar phenotype, while in the presence of blood, the phenotype could be different. The different phenotype could also be related to which hemolysins are being produced by the different strains. The mechanism of heme toxicity in bacteria is not very well understood (333), but *S. aureus* is known to store exogeneously-acquired heme in the cell membrane (334). As membrane proteins, SAOUHSC_01025 and SAOUHSC_01050 could play a role in the regulation of hemolysin production in response to the stores of heme in the membrane.

Figure 34. tn::1025 and tn::1050 mutants have increased hemolysin production. WT, tn::1025, and tn::1050 strains were streaked onto Columbia Blood Agar plates to assess whether they have any kind of hemolysis phenotype. The mutants appeared to secrete more hemolysins than WT as can be seen by the black areas surrounding each colony.

4.3 Expansion of the antibiotic panel and identifying upregulated genes

Advantages of Tn-Seq for predicting mechanism of action

For these experiments, we used the same phage-based transposition system described in the previous chapters (132) with an expanded panel of twenty-five antibiotics in order to develop a method for predicting the mechanism of action of a new antibiotic using Tn-Seq. The phage-based transposition method was originally developed by Meredith and co-workers at Merck Research Labs for the purpose of predicting mechanism of action of new compounds. They showed that they were able to identify the target for many compounds (133, 134). However, this method relied on plating the transposon library on agar plates containing compound at two or three times the MIC and then sequencing the colonies that were able to survive and grow. This method works well, but suffers from two major drawbacks (Figure 35). First, sequencing individual colonies does not give a complete picture of which genes are involved in resistance across the entire genome. Without the significance gained from the millions of reads obtained using NGS, it may be difficult to prioritize and categorize the genes to follow up on. Secondly, many of the new antibiotics we study, especially natural products, are not available in large quantities. The amount of compound needed for use at high concentrations in multiple large agar plates would seriously deplete our supply for most of these compounds and make it impossible to perform many required follow up validation experiments (Figure 35).

Adapting this transposition system for NGS has solved both of these problems. We can treat the transposon library with antibiotic in small (2 milliliter) cultures, which uses only a small amount of compound. Furthermore, because these experiments require very little compound, we can test a few concentrations of the compound to obtain more information about its mechanism of action. It is useful to test multiple concentrations because at lower

concentrations, we can identify the mutants that are more sensitive to the compound, while at higher concentrations, we can identify the mutants that are best able to survive the treatment. Tn-Seq enables the study of compound mechanism of actiomn using a more global, systems-level approach where we can look at transposon insertions in every non-essential gene to determine that gene's relevance to an antibiotic's mechanism of action (Figure 35).
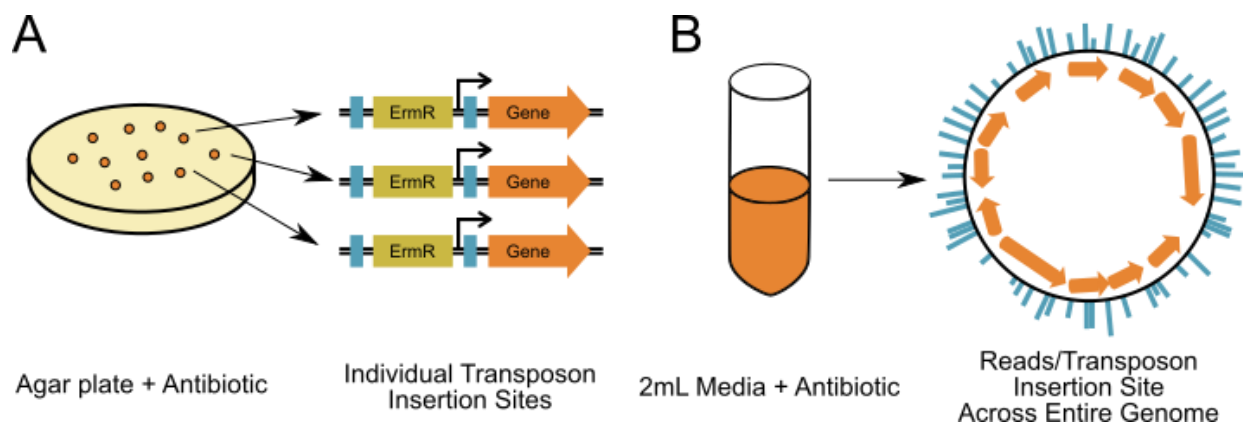
Figure 35. Two approaches for mechanism of action prediction using transposon libraries. (A) The first method relies on selection of a transposon library on an agar plate containing a high concentration of antibiotic. Then, transposon insertion sites are sequenced one colony at a time. Potential targets can be nominated by identifying genes upregulated by the outward-facing promoter of the transposon. This approach has been very successful, but requires a lot of compound and sequencing individual colonies can be time intensive. **(B)** With NGS, the transposon library can be grown in small (2mL) culture volumes and transposon insertion sites can be sequenced using NGS. This allows us to simultaneously assess the fitness of every gene in the genome. A small culture volume allows us to do more experiments with different compound concentrations to identify resistance factors as well as upregulated genes.

Treatment of the transposon library with antibiotic

We have treated the HG003 transposon library (described in Chapters 2) with at least two different concentrations of 25 different antibiotics in 2 milliliters of cation-adjusted TSB (oxacillin and bacitracin were treated in MHB). These antibiotics included many different functional and chemical classes, and a full list can be viewed in Table 8. Concentrations were chosen by identifying at least one antibiotic concentration that caused a moderate delay (approximately three to five hour) and at least one concentration that resulted in an extreme delay (approximately twenty hour) in the time it took for the transposon library to reach stationary phase. Then we collected the cells, extracted the DNA, and prepared the DNA for NGS as previously-described (132). Lower concentrations allow us to identify transposon mutations that sensitize the cell to the antibiotic. At high concentrations, the delay in growth is due to the fact that few mutants are able to grow. This allows us to identify mutants that confer an increase in resistance to the antibiotic.

Table 8. Antibiotics used to treat the transposon library

Mechanism of Action	Antibiotic	Target
Peptidoglycan (PG) Synthesis Inhibition	Oxacillin	Pencillin-binding proteins (Transpeptidase domain)
	Cefaclor	Pencillin-binding proteins (Transpeptidase domain)
	Moenomycin	Pencillin-binding proteins (Glycosyltransferase domain)
	Fosfomycin	MurA
	Cycloserine	Alr, Ddl
	DMPI	MurJ
	CDFI	MurJ
	Vancomycin	Lipid II
	Ramoplanin	Lipid II
	Lysobactin	Lipid I, Lipid II, LipidII$_A$^{WTA}
	Bacitracin	Undecaprenyl pyrophosphate
PG/Wall Teichoic Acid Synthesis Inhibition	Targocil	TarGH
	WTI11	TarGH
Protein Synthesis	Linezolid	50S ribosomal subunit
	Chloramphenicol	50S ribosomal subunit
	Tetracycline	30S ribosomal subunit
	Mupirocin	Isoleucyl-tRNA synthetase
DNA Synthesis	Moxifloxacin	GyrA, Topoisomerases
	Ciprofloxacin	GyrA, Topoisomerases
	Novobiocin	GyrB
	Trimethoprim	Dihydrofolate reductase
	Sulfamethoxazole	Dihydropterate synthetase
RNA Synthesis	Rifampicin	RNA polymerase
Cell Membrane Disruption	Polymixin	Membrane lipids
	Daptomycin	Membrane lipids

Method for identifying genes upregulated by transposon insertion

Because the phage-based transposon method is highly-efficient, it allows for the multiplexing of multiple transposon donor constructs together when creating the transposon library (See Chapter 2) (132, 133). We have taken advantage of this by creating transposon constructs containing outward-facing promoters of different strengths. These regulatory elements allow us to upregulate as well as inactivate any gene in the genome, and this function is very useful for studying a new antibiotic (Figure 36). Not only will we be able to identify resistance mechanisms due to upregulation of genes, but because target-upregulation is a common mechanism of antibiotic resistance, we may also be able to determine the target of a new antibiotic.

We developed a method for identifying genes that, when upregulated, increase resistance to antibiotics. The transposon library was treated with antibiotics and prepared for sequencing as previously described. However, the procedure for data analysis of the sequencing data differs. The goal of this analysis is to identify not only regions of the genome where there is a significant increase in number of reads, but also to identify regions of the genome where there is a preference for orientation of the transposon insertion. Transposons can insert into a TA site with the outward-facing promoter facing either direction, and if upregulation of a gene causes an increase in fitness, there should be a preference for insertion in one direction over the other where the preferential direction corresponds with the same direction as the native promoter of that gene (Figure 36).
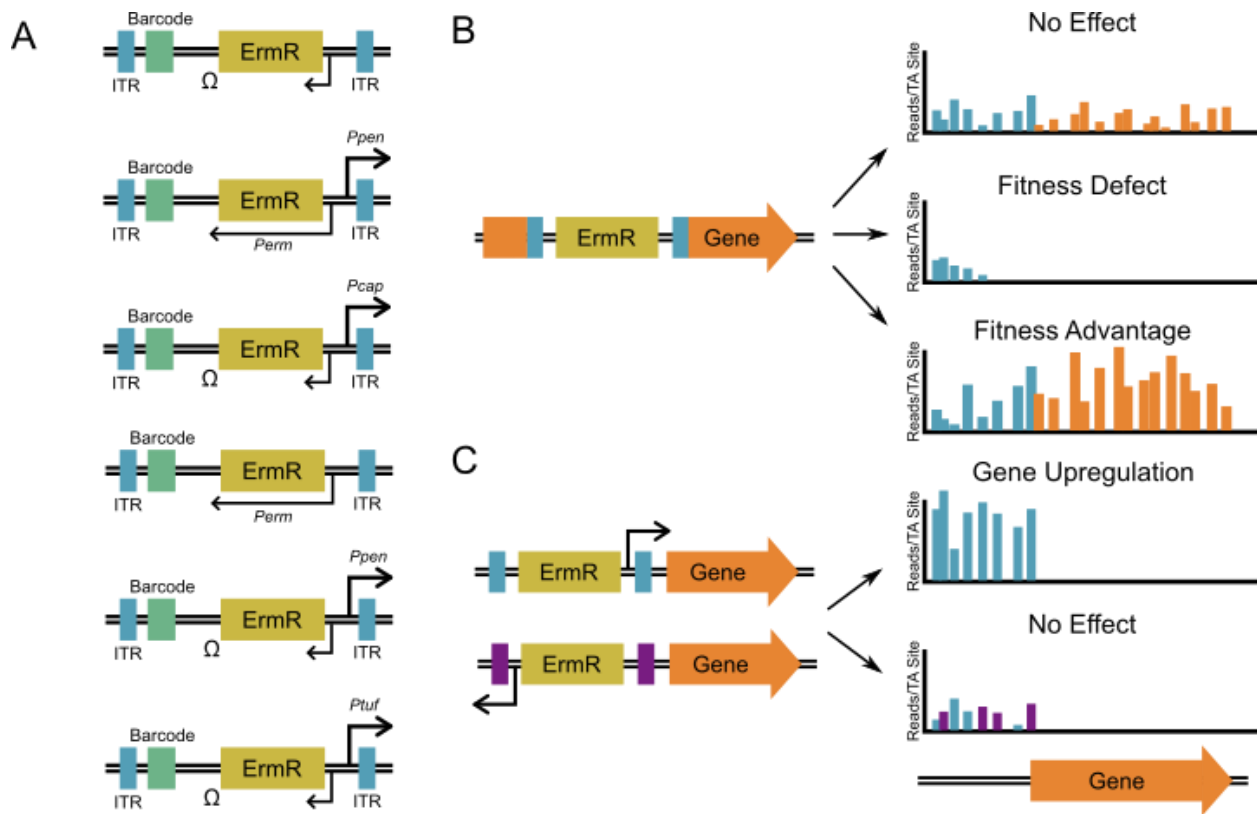
Figure 36. Different transposon constructs can confer different phenotypes. (A) This

transposon library has six transposon constructs multiplexed together (132). These include

(from top to bottom): Blunt construct with no outward facing promoter, Dual construct with P_{pen}

promoter and no transcriptional terminator after the erythromycin resistance gene, Cap

construct with P_{cap} promoter, Erm construct which is the same as the Blunt construct except

without the transcriptional terminator after the erythromycin resistance gene, Pen construct with

P_{pen} promoter, and Tuf construct with P_{tuf} promoter. These outward facing promoters drive the

expression of proximal genes to different extents; in order of increasing strength, P_{erm}, $P_{pen,}$ $P_{cap,}$

and $P_{tuf.}$ **(B)** We can identify different kinds of phenotypes using these constructs. If a

transposon inserts into the middle of a gene, it inactivates that gene. That may have no effect

on the fitness of the mutant, or depending on the condition of interest, the transposon mutant

may confer a fitness defect (fewer number of reads mapping to the gene) or a fitness advantage

(greater number of reads mapping to the gene). **Continued page 117.**

Figure 36 continued. (C) If a transposon construct with an outward-facing promoter inserts ahead of a gene, depending on the orientation of insertion, it may upregulate the proximal or another nearby gene. If the upregulation confers an increase in fitness, more reads will map to that area and there will be a preference for the orientation of transposon insertion. Furthermore, if upregulation of a gene confers a fitness advantage, there will also be few to no reads mapping within that gene because inactivation of that gene usually confers a fitness defect in that condition.

The automated analysis we have developed performs the following operations (Figure 37): 1. It separates the sequencing data into reads mapping to the plus and minus strands for both the control and the antibiotic treatment conditions; 2. it normalizes the data using the simulation based sampling, with the plus strand reads for the treatment condition normalized to the plus strand reads for the control, and the minus strand reads for the treatment normalized to the minus strand reads for the control; 3. because we care about insertions in non-coding regions as well as genes, data is classified into 270bp "windows," with each given a unique numerical name (this window size was chosen because it is ~1/10000th of the length of the genome and should on average contain ~27 TA sites, a window large enough to identify changes due to mutations affecting fitness and small enough to get a fine-grained view of fitness across the genome); 4. it identifies the mean and standard deviation of the number of reads per TA site and the ratio of plus strand to minus strand reads for each TA site across the genome; 5. it identifies TA sites where the number of reads is X standard deviations away from the mean in antibiotic treatment but not the control to identify TA sites where transposon insertion confers a fitness advantage; 6. it identifies TA sites where the number of reads in one strand is X standard deviations away from the mean but not in the other, suggesting that there is a preference for orientation of transposon insertion. X can be increased or decreased until 100-

117

200 TA-sites remain, depending on how stringent of an analysis is desired. In our experience, increasing the stringency to approximately 100 TA sites allows us to be much more confident in our results. Decreasing the stringency can make analysis difficult because distinguishing truly upregulated genes from the background noise becomes more challenging. Here, we have used relatively stringent cutoffs.

The rest of the analysis is best done manually in order to retain all interesting genes (Figure 37). We identify windows where there are three or more TA sites that meet the cutoffs described above. Then, we determine where in the genome that window is and which direction the outward-facing transposon promoter is facing. Based on this information, we can nominate candidate genes that could be upregulated by the transposon insertions in this window. These promoters have the ability to act on genes not immediately proximal to the transposon insertion (133), so for this work, we looked for a gene within 2kb facing the appropriate direction. Then, we returned to the raw reads per TA site data to confirm that the candidate gene is upregulated. A gene whose upregulation increases the cell's fitness with antibiotic treatment should have few or no reads mapping to it because inactivating the gene is likely to cause a decrease in fitness in that condition. Furthermore, in non-coding regions and in non-essential genes ahead of the candidate gene, we should see an increase in number of reads and a strong preference for transposon orientation (Figure 37).

Figure 37. (See page 120) Automated upregulation analysis indentifies signatures of upregulation and nominates candidate upregulated genes. 0. The starting point for this analysis uses the raw reads per transposon insertion site data (red bars). **1**. The analysis begins by separating the sequencing data into reads mapping to the plus (purple bars) and minus (blue bars) strands for both the control and the antibiotic treatment conditions. **2.** Then, it normalizes the data using the simulation based sampling. **3.** Each read is then mapped to unique 270bp "windows" (numbered boxes below chart) with a unique numerical identifier. **4.** Next, the mean number of reads per TA site and the standard deviation of that value across the genome is calculated. We also calculate the mean and standard deviation of the ratio of plus strand to minus strand reads for each TA site. **5.** We identify the TA sites with an increase in reads compared to the average (X standard deviations from the mean, where X is empirically-determined). Here, TA sites with such an increase are marked with an asterisk. **6.** Then, of those TA sites with an increase in reads, we identify those with a preference for orientation of transposon insertion, by identifying an increase or decrease in the ratio of plus to minus strand reads compared to the average (X standard deviations away from the mean, whre X is empirically-determined). These TA sites are shown with a large red asterisk in contrast to the small black ones. **7.** To identify signatures of upregulation, we identify windows where there are three or more TA sites that meet the cutoffs described above or adjacent windows with at least one TA site meeting the described cutoff. **8.** Finally, we map the hit window back to the genome, and nominate candidate genes based on the location of the signature of upregulation and the orientation of that signature. In this example, the green Gene 2 is immediately proximal to the signature of upregulation and has no reads mapping to it, suggesting that it is essential in this condition and making it a likely candidate for upregulation.

Continued. This analysis can, in an unbiased way, distinguish between an increase in fitness due to upregulation of a gene (Gene 2) and an increase in reads that is due to inactivation of a gene (Gene 2).

Figure 37 Continued.

0. Reads mapped to TA sites
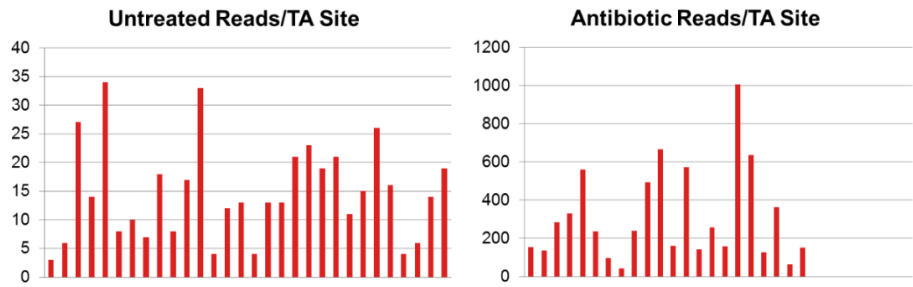
1. Split by reads mapping to plus and minus DNA strands
2. Normalization
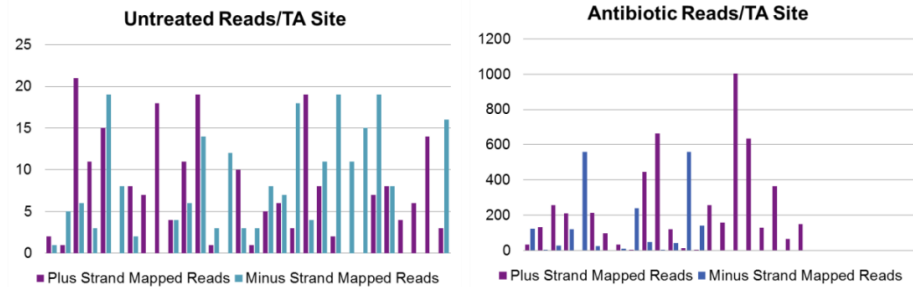
3. Map TA site to 270bp windows
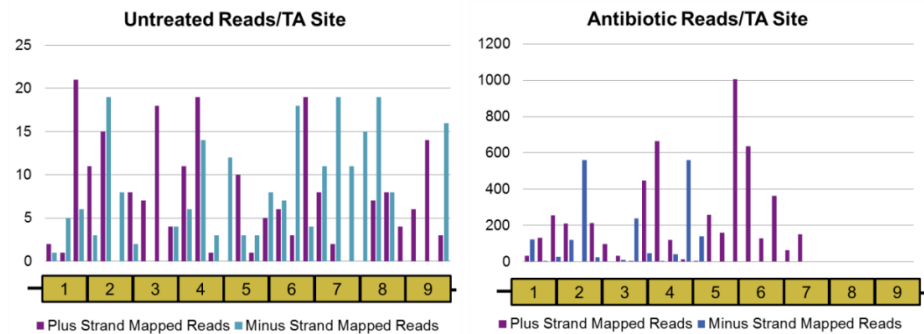
4. Calculate mean reads/TA site
5. Identify TA sites with increase in number of reads
6. Identify TA sites with increase in reads mapping to only one strand
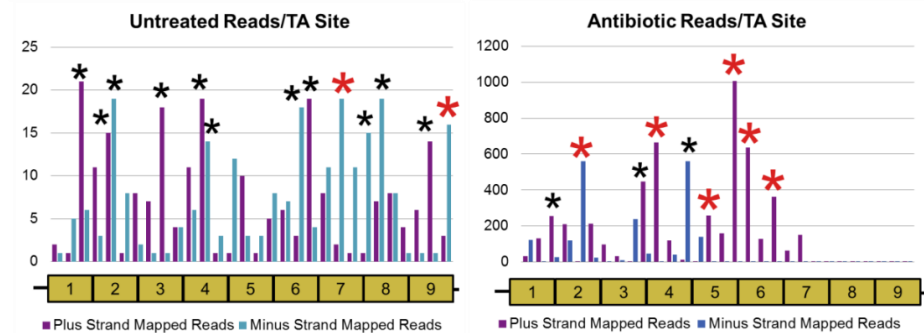
7. Identify adjacent windows with signature of upregulation
8. Based on orientation of transposon insertion, identify nearby genes that could be upregulated

121
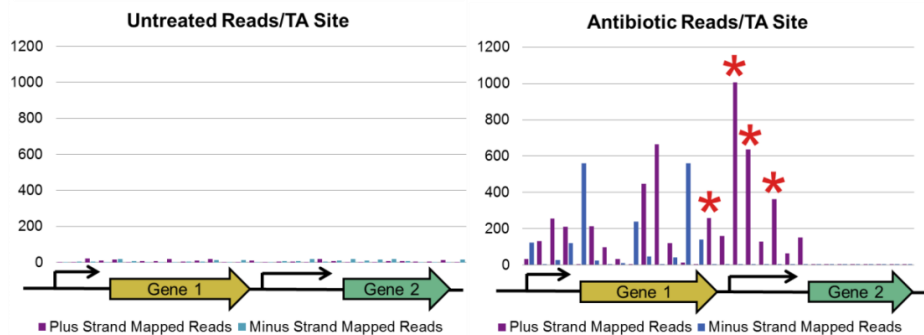
Genes upregulated by transposon insertion

We have validated this method by performing it with every concentration of every antibiotic in Table 8. This method generally works better when the antibiotic concentration is close to or above the MIC. At high antibiotic concentrations, more transposon mutants are sensitive, so the background level of reads per transposon insertion site is low and it is easy to identify regions where transposon insertion may be upregulating genes.

We identified upregulated genes for 11 of the 25 antibiotics, with 18 total genes upregulated (Table 9), including both essential and non-essential genes (Figure 38). Of these 18 genes, half were target upregulation or target-modifying genes, three were hypothetical genes, and the rest were other annotated genes that, when upregulated, confer resistance through both known and unknown mechanisms (Table 9). The types of known resistance mechanisms included upregulation of efflux pumps (ex: *norA* with ciprofloxacin treatment (298)), antibiotic modification genes (ex: *fosB* with fosfomycin treatment (8)), target modification genes (ex: *uppP* with bacitracin treatment (335)), other clinical resistance mechanisms (ex: *mprF* with daptomycin treatment (52, 55, 279, 336)), and target upregulation (ex: *alr* and *ddl* with cycloserine treatment (337)) (Table 9). Of the novel resistance mechanisms we identified, some were known resistance mechanisms to other antibiotics but had never been seen for this antibiotic (ex: *uppP* with moenomycin treatment), some were known genes, but they had never been shown to confer resistance when upregulated (ex: *recO* with moxifloxacin treatment (324)), and some were completely hypothetical genes that have never been studied before (ex: *SAOUHSC_02149* with daptomycin treatment) (Table 9).

Table 9. Genes upregulated by transposon insertion		
Antibiotic	**Upregulated Gene**	**Type of Resistance**
Bacitracin	uppP	Target modification
Ciprofloxacin	norA	Antibiotic efflux
Moxifloxacin	recO	Novel mechanism
Fosfomycin	fosB	Antibiotic modification
	murA	Target upregulation
	murA	Target upregulation
Daptomycin	mprF	Known mechanism
	SAOUHSC_00969	Novel mechanism
	SAOUHSC_02149	Novel mechanism
	SAOUHSC_02164	Novel mechanism
DMPI	murJ	Target upregulation
CDFI	murJ	Target upregulation
Moenomycin	uppP	Novel mechanism
Cycloserine	alR	Target upregulation
	ddL	Target upregulation
Novobiocin	gyrB	Target upregulation
Cefaclor	tarGH	Novel mechanism
	vraRS	Cell wall stress response
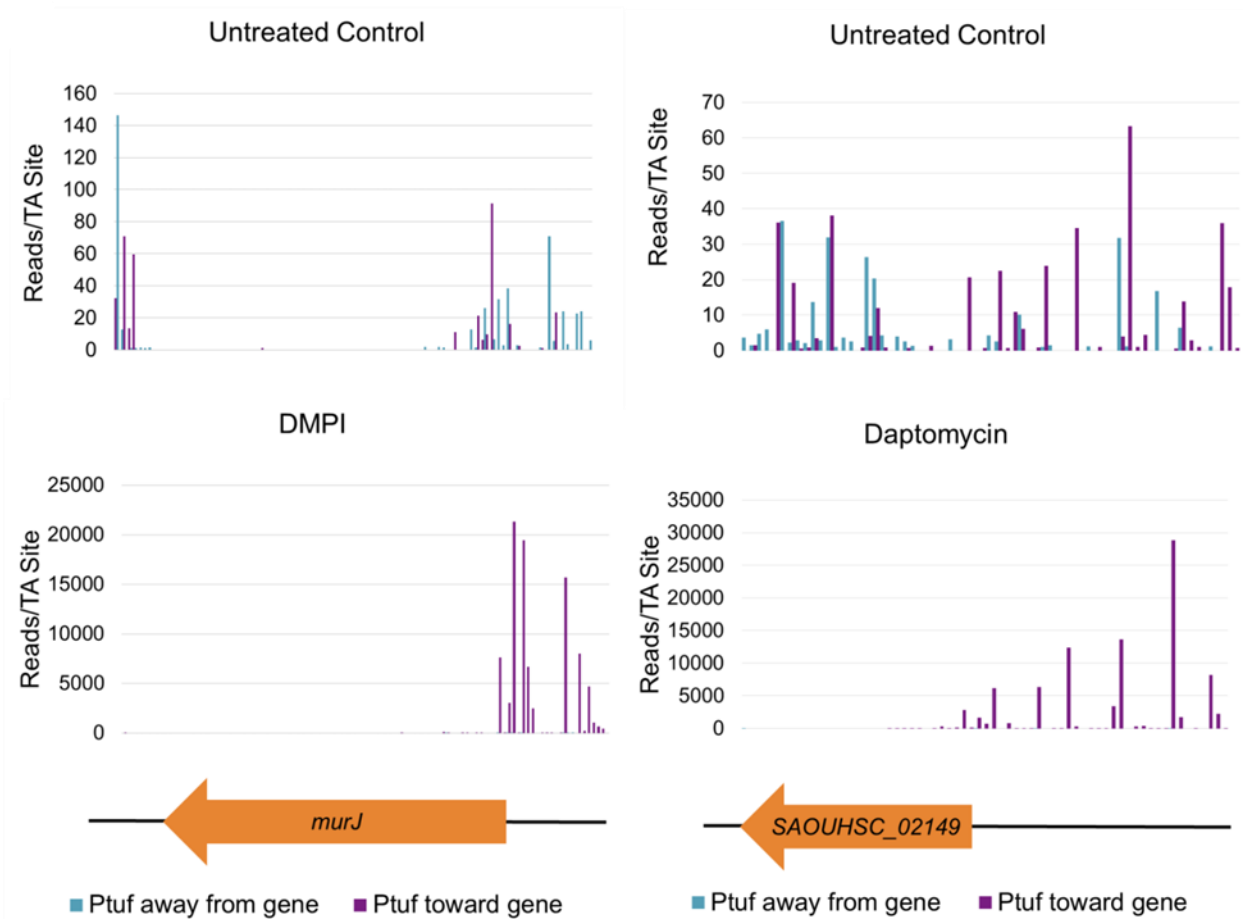
Figure 38. Both essential and non-essential genes are found to be upregulated by transposon insertion. We manually confirm that our method for identifying these genes is functioning correctly by plotting the number of reads/TA site mapping to each strand for both the control and antibiotic-treated data. Here, we show the results for an essential gene, *murJ,* and a non-essential hypothetical gene, *SAOUHSC_02149* using the construct carrying the P_{tuf} promoter. In the untreated control datasets, reads map to both strands of the genome meaning that transposons insert with the outward-facing promoter facing both toward and away from the gene. No reads map to *murJ* itself because it is essential (338). However, when you treat with an antibiotic, not only is there a 100-1000 fold increase in number of reads that map to the promoter region, but there is a complete preference for the orientation of the transposon such that the P_{tuf} promoter faces toward the gene.

There are several possible explanations why we were not able to identify upregulated genes for every antibiotic. First of all, five of the antibiotics tested (vancomycin, ramoplanin, bacitracin, daptomycin, and polymixin) do not bind to and inhibit enzymes (15, 18, 19, 339-341), making identification of target upregulation for these antibiotics impossible. Secondly, this method relies on growing the library in liquid culture instead of on plates. Growing in liquid culture results in more competition between every mutant in the library. On a plate, a mutant may confer an increase in resistance, but if it does not also confer a high level of fitness compared to every other mutant in the library, when grown in liquid culture, it may be out-competed by another mutant with a higher fitness. Some genes (often encoding membrane proteins) are known to have a fitness defect when upregulated (342, 343), and occasionally, upregulation of a target can actually sensitize the cell to antibiotic treatment (135). Therefore, it may be difficult to identify upregulated genes conferring an increase in resistance to an antibiotic using this method unless those genes also do not confer any kind of fitness defect.

Validation of novel mechanisms of resistance

We validated a selection of the known and novel mechanisms of resistance. To do this, we placed the candidate gene of interest on the pLOW plasmid under the control of a strong constitutive promoter in the HG003 strain. We compared the growth of this strain with WT as well as a strain that has a transposon insertion inactivating the gene of interest. This strain was obtained from the Nebraska transposon library (130) and transduced into the HG003 background. We chose to validate some of the novel genes identified above: *uppP* for moenomycin and the three hypothetical genes for daptomycin: *SAOUHSC_00969, SAOUHSC_02149,* and *SAOUHSC_02164*. We used *uppP* in our validation strategy because its known resistance activity with bacitracin will be a control for this validation strategy, and we chose to validate the novel genes identified with daptomycin because daptomycin is a relatively new antibiotic with a significant amount of success in the clinic. *SAOUHSC_00969,*

SAOUHSC_02149, and *SAOUHSC_02164* are all very small with 107, 175, and 60 amino acids, respectively. Only *SAOUHSC_02149* has any conserved domains, with a predicted bacterial Pleckstrin homology (bPH) domain at its N-terminus. These domains in eukaryotes bind phosphatidylinositol and are involved in recruiting proteins to membrane regions. All three of these proteins are predicted to have membrane spanning regions. The membrane spanning regions for SAOUHSC_00969 and SAOUHSC_02149 are near the N-terminus, while the membrane spanning region for SAOUHSC_02164 is in the C-terminal half. Other small peptides have been shown to be upregulated during cell wall stress (344), suggesting that daptomycin may be inducing cell wall stress along with membrane damage. However, these genes were not identified as hits with any other cell envelope targeting antibiotic, and upregulation of *SAOUHSC_02149* did not confer an increase in resistance to any other antibiotic (Figure 39), suggesting that this phenotype is specific to daptomycin for at least one of these genes.
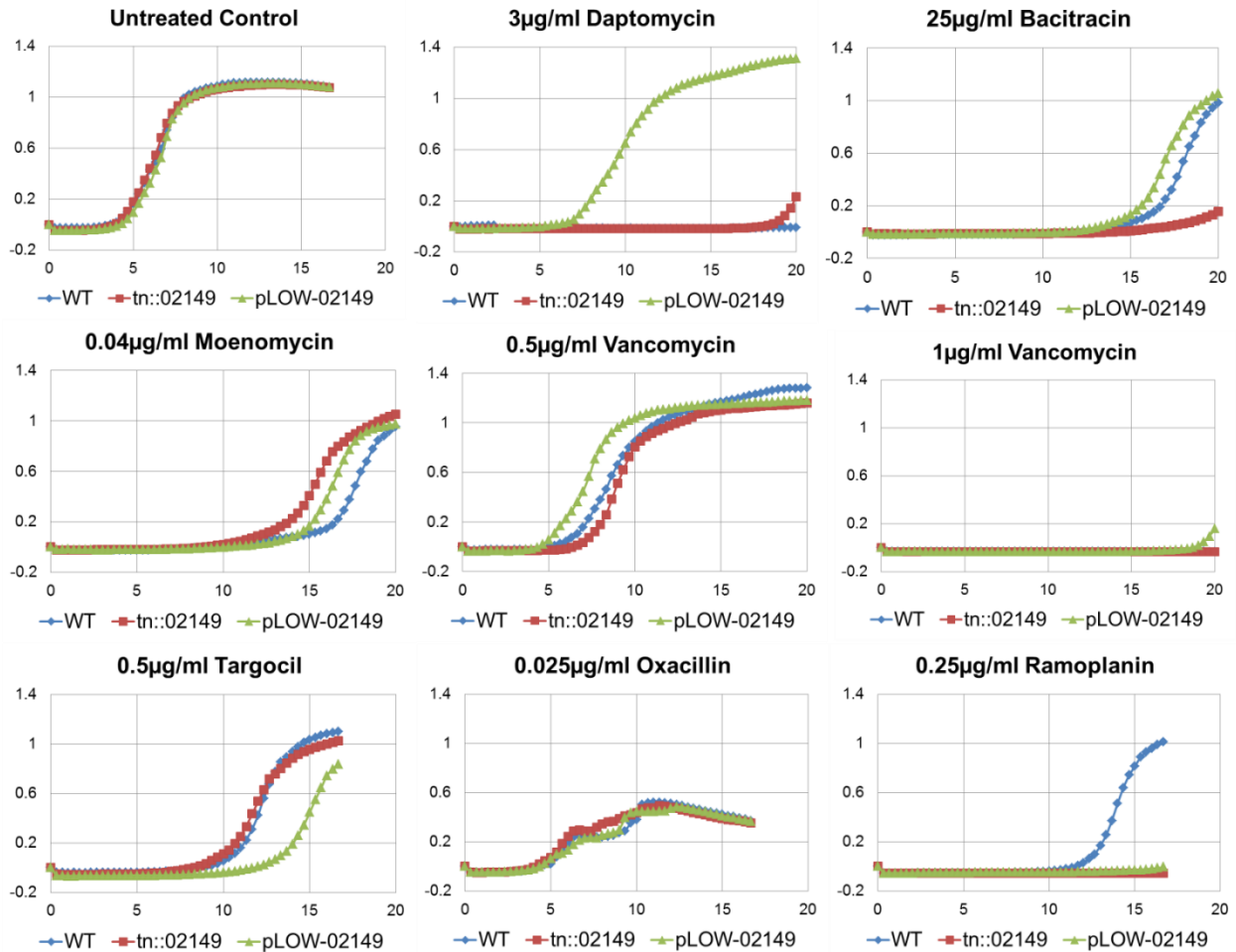
Figure 39. Upregulation of *SAOUHSC_02149* only protects from daptomycin. WT, the transposon inactivation mutant of *SAOUHSC_02149* (tn::02149), and a constitutively-upregulated version in the pLOW plasmid (pLOW-02149) were grown without antibiotic and with daptomycin, bacitracin, moenomycin, vancomycin, targocil, oxacillin, and ramoplanin in a 96 well plate to see if upregulating this gene protected from any other type of cell wall stress. All of these antibiotics inhibit cell wall synthesis. There was no significant change in growth compared to WT for pLOW-02149 in the presence of bacitracin, moenomycin, or oxacillin. **Continued page 128.**

Figure 39 continued. At 0.5µg/ml vancomycin, it seemed that pLOW-02149 might have better growth than WT, while tn::02149 had poorer growth, but at the next higher concentration tested, 1µg/ml vancomycin, all strains were dead suggesting that *SAOUHSC_02149* does not exert much of a protective effect. Furthermore, pLOW-02149 does not protect from targocil and ramoplanin treatment, and in fact, this strain may be slightly sensitive to these antibiotics. .

Growth curves were used for validation of these unknown genes. Cells were grown in 96-well plates for 16-18 hours at 37°C with and without antibiotics. Though there was no difference in fitness in the absence of antibiotic treatment, in the presence of antibiotics, the transposon mutants grew more slowly or did not grow at all, while the upregulated genes grew to stationary phase before the WT cells did (Figure 40). However, some upregulation constructs produced a higher increase in resistance than others. Upregulation of *uppP* increases resistance to bacitracin at least 32 fold while the other genes produce at most a more modest 2-4 fold increase in MIC. All the results tested do validate to some extent, proving that this method works well for identifying genes that when upregulated by transposon insertion confer a fitness advantage.
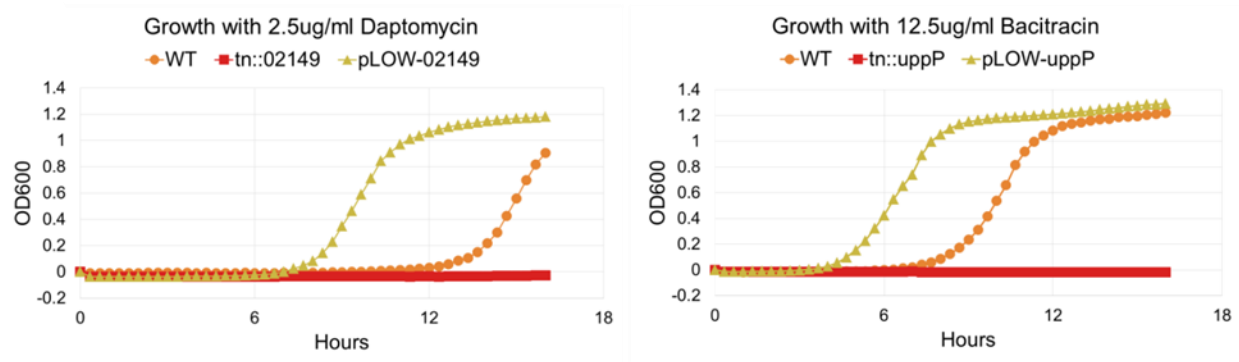
Figure 40. Genes upregulated by transposon insertion confer increased antibiotic resistance. A selection of the validation performed is shown. We validated some of these hits by placing the gene of interest onto a plasmid that constitutively upregulates the gene. Then we compared the growth of these strains in the presence of antibiotic with both WT and a strain with the same gene knocked out. Here, we show a novel mechanism of resistance to daptomycin (upregulation of *SAOUHSC_02149*) and as a control, a known mechanism of resistance to bacitracin (upregulation of *uppP*) (335). As expected, the knockout is more sensitive to antibiotic treatment than WT, and the upregulated strain is more resistant than WT.

4.4 Applying machine learning algorithms to Tn-Seq data to predict antibiotic mechanism

Strategy for predicting mechanism of action

Though the method for identifying upregulated genes works well, we did not observe target upregulation for every antibiotic. Therefore, if we want to use this platform to determine antibiotic mechanism of action, we need to use another strategy. The strategy we have chosen to use takes advantage of an observation we made about vancomycin and ramoplanin. Vancomycin and ramoplanin have the same target, the peptidoglycan precursor, lipid II, but they bind to different parts of the molecule (16, 339). While vancomycin binds to the D-Ala-D-Ala moiety, ramoplanin binds to the sugar phosphate head. Because they have the same target molecule, the set of transposon mutants that can confer resistance and sensitivity to these antibiotics are very similar. If we look at every gene in the genome for each antibiotic and whether transposon insertions in that gene confer resistance of sensitivity, we can obtain a unique "resistance factor fingerprint" for each antibiotic. We noticed that the fingerprints for ramoplanin and vancomycin were almost identical, and we wondered whether this was true for the other antibiotics in our panel that had the same target or that targeted the same pathway. A cursory examination of the data revealed that this fingerprint is very similar for antibiotics with similar mechanisms of action such as trimethoprim and sulfamethoxazole, ciprofloxacin and moxifloxacin, as well as DMPI and CDFI.

For a more quantitative comparison of the data, we used hierarchical clustering (Figure 41). We chose to compare only one concentration for each antibiotic, always using sequencing data from the traditional "Blunt" transposon construct. We chose these concentrations by identifying the concentrations that had close to or greater than 1 million reads with a significant change in number of TA sites with insertions compared to the control. We looked for conditions where the antibiotic-treated condition only had about 40 - 60% of the TA insertion sites as the

control. Then, to compare data from different antibiotics tested, we had to put all the data on the same scale. We used the same method described earlier in this chapter to calculate a "fitness value" for each gene with each antibiotic treatment. Once we had the data on the same scale, we used hierarchical clustering to assess the similarities between antibiotics. This allows us to visualize the fingerprints and how similar they are for different antibiotics (Figure 41). Antibiotics with the same target such as vancomycin and ramoplanin (16, 339) or DMPI and CDFI (345) cluster together. In addition, there is a dramatic difference between some of the peptidoglycan synthesis inhibitors and all the others, highlighting the fact that the genes required for resistance to this class of antibiotics are very different than non-cell envelope targeting antibiotics.
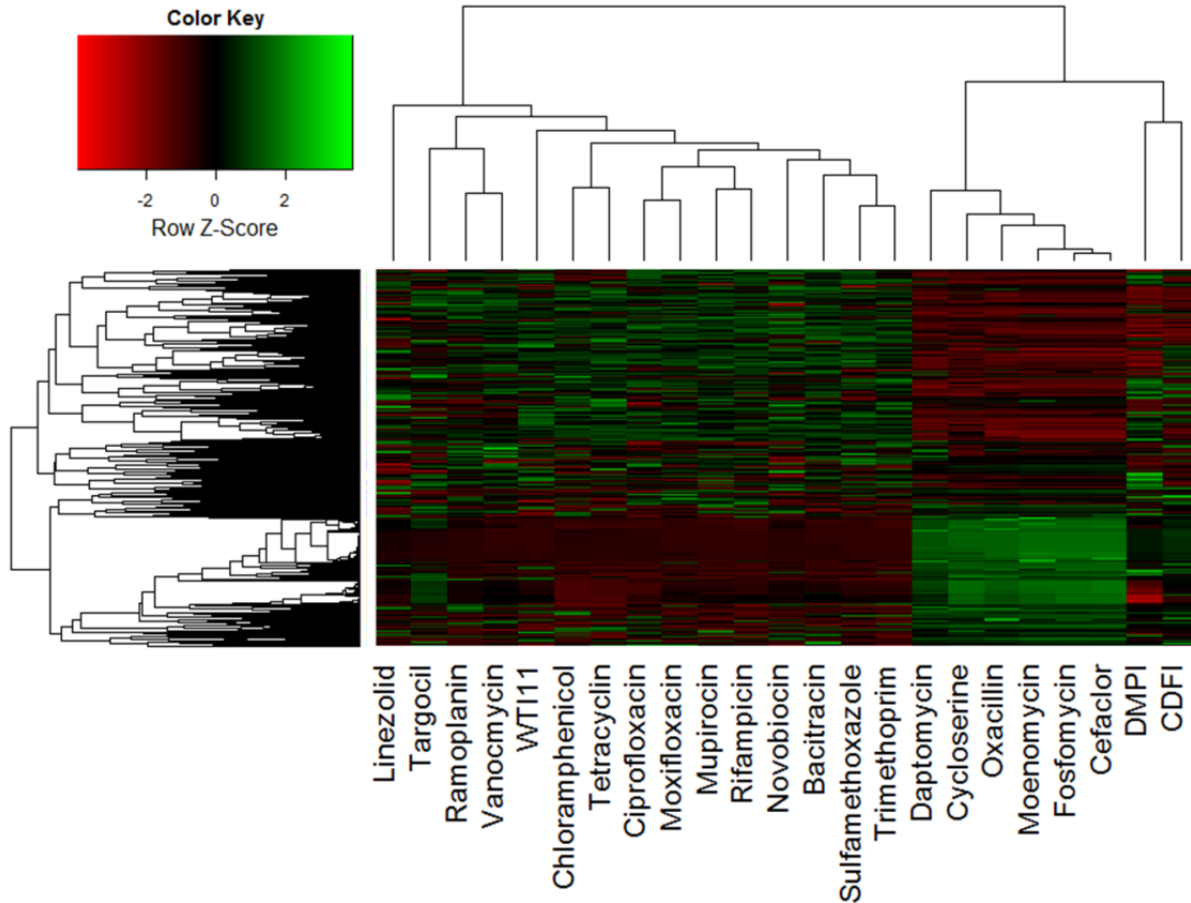
Figure 41. Some antibiotics cluster by mechanism of action using hierarchical clustering. Hierarchical clustering is a good way to help visualize the differences between antibiotic datasets. Here, we compare the fitness values for a subset of genes in the genome for a selection of antibiotic treatments. Bright red genes correspond to genes where transposon insertion confers a fitness defect, while bright green genes have a fitness advantage when inactivated with a transposon. The most obvious cluster is that consisting of cycloserine, oxacillin, moenomycin, fosfomycin, and cefaclor (all peptidoglycan synthesis inhibitors), but other compounds with the same target also happen to cluster together including ramoplanin/vancomycin, ciprofloxacin/moxifloxacin, and DMPI/CDFI. **Continued page 133.**

Figure 41 Continued. Though this method for classifying antibiotics can not be used to predict the mechanism of action of a new antibiotic, it allows us to visualize some of the similarities and differences in resistance factors between different antibiotic treatments.

Using the K-nearest neighbors algorithm as an antibiotic classifier

We can observe differences in patterns with the heatmap output by hierarchical clustering, but we needed a better method for predicting mechanism of action using the resistance factor fingerprint of a new antibiotic with an unknown mechanism of action. Earlier in this chapter, we took advantage of the utility of machine learning algorithms to identify genes that when inactivated had the same pattern of resistance and sensitivity across a small panel of antibiotics, which helped us predict the pathways a hypothetical gene may be a part of. We can use a similar approach for predicting the mechanism of an unknown antibiotic. Using machine learning classifier algorithms, we can take advantage of information we already know about the mechanisms of action of our large panel of antibiotics in order to predict the target of a new antibiotic. After much experimentation with different classifiers, we again chose the K-nearest neighbors algorithm as implemented by the sci-kit learn Python library (318), the same algorithm earlier used to learn more about hypothetical genes. We can use our panel of antibiotics as the training data for this algorithm, telling the algorithm which class of antibiotics each known antibiotic belongs to. Then, when testing a new antibiotic, we can determine which class of antibiotics, a new antibiotic is most similar to (Figure 42).
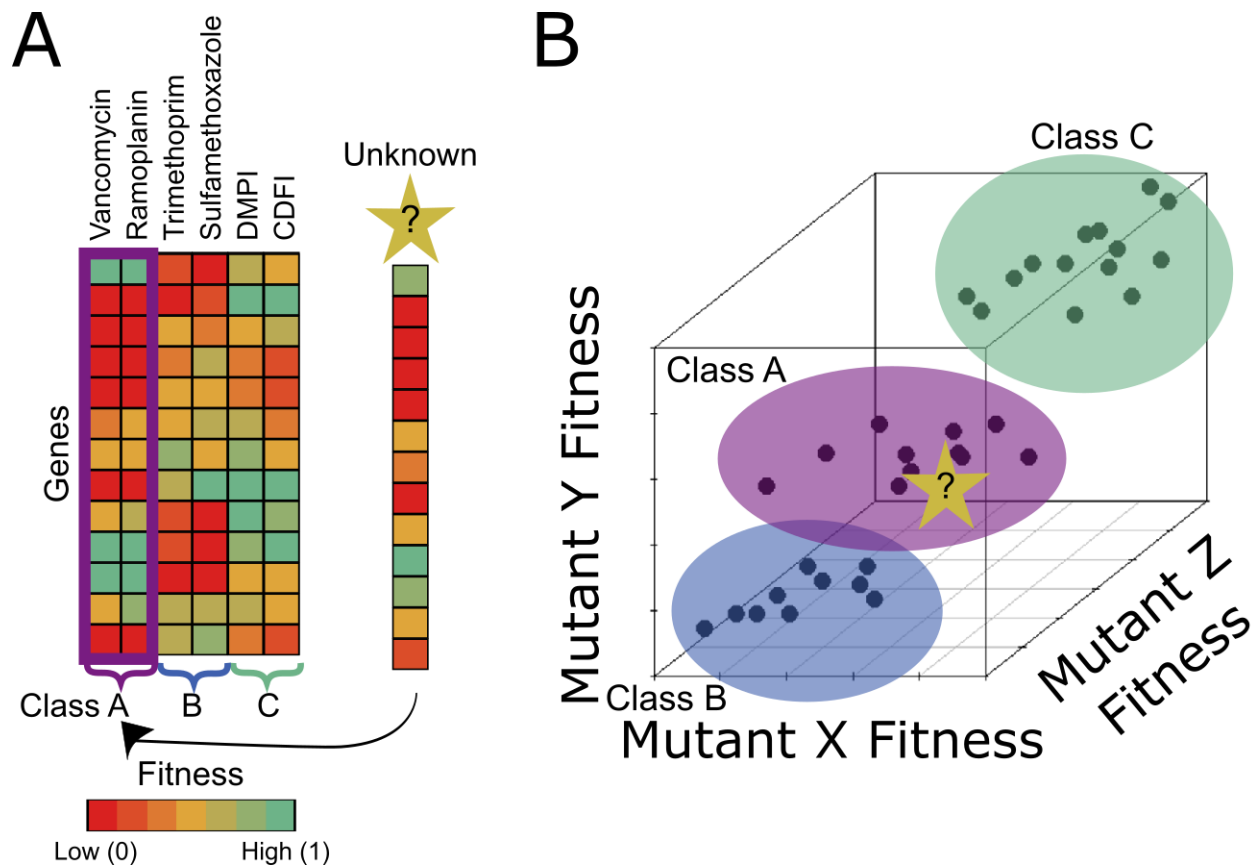
Figure 42. Machine learning strategy for mechanism of action prediction. (A) We observed that the transposon mutations conferring resistance and sensitivity to antibiotics were very similar for antibiotics with the same target or with a target in the same pathway. We wanted to use this resistance factor fingerprint to predict the mechanism for a new antibiotic with an unknown target by comparing the resistance factor fingerprint of the new antibiotic to a curated panel of antibiotics with known mechanisms of action. **(B)** To do this, we use the K-nearest neighbors classifyin algorithm. A hypothetical example of the supervised K-nearest neighbors classifier (318) is shown using only three genes. This algorithm works by using a training dataset of antibiotics with known classes. Then, it takes an unknown antibiotic and uses the distance of that antibiotic to its closest X-number of neighbors to predict which class that antibiotic belongs to. In this case, the novel antibiotic is most similar to class A antibiotics suggesting that it may bind to Lipid II as vancomycin and ramoplanin do.

Additionally, this algorithm gives us the option of performing either supervised or unsupervised learning, depending on the goals of the analysis. Unsupervised learning takes the data for each antibiotic, and tries its best to compare them without any input from the user. In that sense, it is similar to hierarchical clustering, and as expected, the data from performing unsupervised learning resembles the results from the hierarchical clustering. When we use this algorithm for performing supervised clustering, we can tell it which class of antibiotics each antibiotic dataset belongs to. Then, when we predict the mechanism of a new antibiotic, the algorithm outputs which class of antibiotics the new one is most similar to, in contrast with unsupervised learning, which returns the most similar antibiotics.

The next question was how to classify the antibiotics used in the training set. The most general classifier would be one that distinguished between cell-envelope targeting and non-cell envelope targeting antibiotics. However, this is not very useful for predicting the target of an antibiotic. On the other hand, going too detailed with our classifier and classifying by target itself would also not work well for a dataset of our size because our training set would only consist of one or two antibiotics per category. With a larger training set consisting of many antibiotics for each target of interest, we would likely be able to more precisely classify these antibiotics, but as a proof of principle project, we chose somewhat broader classes which reflect our smaller dataset (Table 10). These classes were chosen by examining the results of the hierarchical clustering and unsupervised learning analyses. For some classes of antibiotics where we had multiple antibiotics with the same target in our panel, we were able to go down to the target level as a classifier (ex: vancomycin and ramoplanin can be reliably classified as lipid II binding antibiotics). However, for others, it was more difficult to identify a group with similarities, so other antibiotics are classified into broader categories (ex: rifamycin and mupirocin were classified together as non-cell envelope targeting antibiotics). Unfortunately, a few antibiotics were tested

at too low of a concentration, and so those antibiotics were not able to be accurately placed into

a classifier (Table 10).

Table 10. Classes of antibiotics for supervised machine learning	
Target pathways	**Antibiotic**
Peptidoglycan (PG) Synthesis	Oxacillin
	Cefaclor
	Moenomycin
	Fosfomycin
	Cycloserine
	WTI11 (via Lipid II Depletion)
PG Synthesis via Lipid II Flippase	DMPI
	CDFI
PG Synthesis via Lipid II Binding	Vancomycin
	Ramoplanin
Membrane Disruption	Polymixin
	Daptomycin
Protein Synthesis	Chloramphenicol
	Tetracycline
	Linezolid
DNA Gyrase	Moxifloxacin
	Ciprofloxacin
Folate Synthesis	Trimethoprim
	Sulfamethoxazole
Other non cell envelope target	Rifampicin
	Mupirocin
Unclassified	Targocil
	Bacitracin
	Lysobactin
	Novobiocin

Training the classifier algorithm

Various parameters were modified for optimizing this algorithm including the number of nearest neighbors, the distance metric, and the number of *S. aureus* genes included in the analysis. To assess the progress of the optimization, we used the unsupervised learning algorithm and four different metrics utilizing some of our known antibiotics as "test unknowns". These four metrics are described in order of increasing stringency. **1) One antibiotic at one concentration should be most similar to the same antibiotic at a different concentration.** For this, we used the 0.16ug/ml Moenomycin concentration which should be most similar to the 0.32ug/ml Moenomycin sample. We want this method to be somewhat insensitive to the concentration at which the library is treated. Of course, if the concentration is too low or too high, no amount of normalization or training will be able to classify an antibiotic correctly, but we would like this algorithm to be able to tolerate small changes in antibiotic concentration**. 2) One compound of a specific chemical class should be most similar to another compound of the same chemical class if they have the same target.** For this metric we used CDFI, predicting it would be most similar to DMPI. These are two antibiotics of the same compound class whose target is the lipid II flippase, MurJ (345). **3) Two antibiotics of different chemical classes, but with the same target should also be most similar to each other.** Depending on an antibiotic's structure, charge, and method of entering the cell, it may elicit different stress responses, causing different transposon mutants to be differentially sensitive between the two antibiotics. However, we would still like to classify antibiotics by their mechanism as opposed to their chemical class. For this comparison, we used targocil and WTI-11, two compounds which target the wall teichoic transporter, TarGH, but kill the cell by depleting the lipid tether also used by peptidoglycan (PG) biosynthesis (324, 346). We found in early testing that the lower concentration of WTI-11 has a similar set of resistance factors as targocil, but the higher concentration appears more similar to peptidoglycan biosynthesis inhibitors. This makes sense

because the highest concentration at which we treated the library with targocil did not exert much selective pressure, and few significant hits were obtained. Therefore, it is not surprising that only the lower WTI-11 concentration matches targocil. The lower concentration of WTI-11 was used here, but the higher one was used for the final clustering. These differences highlight the importance of testing different antibiotic concentrations. **4) Our most stringent metric for training this algorithm was using two antibiotics of different chemical classes with different targets, but whose targets are in the same pathway.** We hoped that we could classify inhibitors into pathways, which would drastically decrease the number of possible targets compared to all the genes in the genome, or even all the genes involved in the cell envelope. For this, we used trimethoprim and sulfamethoxazole which both target different enzymes in the folate pathway (10, 347). We optimized the algorithm such that all four metrics were satisfied. In the end, we chose to use the Minkowsky distance metric, 2 nearest neighbors, and we found that the best classification resulted from only using a subset of genes. For each antibiotic, we identified the top 25% and bottom 25% with the biggest change in fitness values, and we used the unique set of these genes for this analysis, a total of 1614 genes.

Validation of the method

In order to validate the method, instead of only using four antibiotics as we did in the optimization, we used all 25 antibiotics as "test unknowns", and classified them with our optimized algorithm using all the other 24 antibiotics as the training set. The categories we used for this analysis are shown in Table 10. Unclassifiable antibiotics do not cluster well with drugs that are similar to them, likely because the concentration tested was too low.

The results of our validation are shown in Table 11. We defined success in two ways: by whether or not the antibiotic was classified to the correct category and by whether or not we were able to learn about the antibiotic's mechanism of action based on its categorization. The difference between these two is whether or not it was categorized into the "Unclassified"

138

category. Overall, this algorithm classifies the antibiotics as we expected it to 80% of the time.

There is some debate over the mechanism of action of daptomycin, which disrupts the

membrane but also delocalizes some of the PG synthetic enzymes (18, 19). We observe that

the transposon mutations conferring resistance and sensitization to daptomycin more closely

match that of peptidoglycan biosynthesis inhibitors than other membrane disruptors such as

polymixin. This result suggests that its mechanism of killing may be due to its PG synthesis

inhibitory activity as opposed to its membrane disrupting characteristics. If we include

daptomycin in our results as correct, then, we predict mechanism of action accurately 72% of

the time. This is as good or better than other published computational mechanism prediction

methods. The method most similar to ours was published in 2015 by Wildenhain and coworkers.

This method also takes advantage of machine learning to identify similar patterns of chemical-

genetic interactions to identify the target of a drug in yeast (290). Their algorithm, SONAR[G],

correctly predicts the target or a target related gene among the top six gene hits for 16 out of 27

of the compounds tested (290). This corresponds to a 59% success rate. It is likely that with a

larger training data set, our classifier would be able to more accurate identify the mechanism,

and with more training data per antibiotic target, we would be able to more reliably predict the

target of the antibiotic.

Table 11. Mechanism of action predicted by machine learning analysis

Validated Mechanism of Action	Antibiotic	Predicted Target Pathways
Peptidoglycan (PG) Synthesis Inhibition	Oxacillin	PG Synthesis
	Cefaclor	PG Synthesis
	Moenomycin	PG Synthesis
	Fosfomycin	PG Synthesis
	Cycloserine	PG Synthesis
	DMPI	PG Synthesis via Lipid II Flippase
	CDFI	PG Synthesis via Lipid II Flippase
	Vancomycin	PG Synthesis via Lipid II Binding
	Ramoplanin	PG Synthesis via Lipid II Binding
	Lysobactin	Folate Synthesis
	Bacitracin	Unclassified
PG/Wall Teichoic Acid Synthesis Inhibition	Targocil	PG Synthesis via Lipid II Binding
	WTI11	PG Synthesis
Protein Synthesis	Linezolid	Protein Synthesis
	Chloramphenicol	Protein Synthesis
	Tetracycline	Protein Synthesis
	Mupirocin	Folate Synthesis
DNA Synthesis	Moxifloxacin	DNA Gyrase
	Ciprofloxacin	DNA Gyrase
	Novobiocin	Unclassified
	Trimethoprim	Folate Synthesis
	Sulfamethoxazole	Folate Synthesis
RNA Synthesis	Rifampicin	PG Synthesis via Lipid II Binding
Cell Membrane Disruption	Polymixin	Protein Synthesis
	Daptomycin	PG Synthesis

Table 11. Mechanism of action predicted by machine learning analysis. Column 1 shows the target pathways for the antibiotics in Column 2. Column 3 shows the predicted mechanism of action using our machine learning approach. Cells are colored green (correct prediction), yellow (correct unclassified prediction), and red (incorrect classification).

4.5 Application of this approach to predict the mechanism of action of a new antibiotic

The WAP antibiotics

Next we wanted to test this algorithm using a new compound with an unknown mechanism of action. We have access to a series of compounds extracted from *Myxococcal* fungi that we are working on in collaboration with Rolf Muller's lab. I chose to concentrate on attempting to predict the mechanism of the WAP compound series (WAP1, WAP2, and WAP3) because we may be able to learn more about the target from a series of related compounds than if we just tested one antibiotic . In essence, it gives us more chances to guess the correct mechanism. Furthermore, as a set of related compounds, they should have the same or similar targets. If we see the same mechanism of action as a hit for more than one of the WAP compounds, we can be more confident in the results of that prediction. The WAP compounds are peptide-based antibiotics whose mechanism is unknown. One related compound WAP-8294A2 was in phase I/II clinical trials as of 2011 (348, 349). The mechanism of this related compound is also not known. However, it is known that cardiolipin and phosphatidylglycerol antagonizes its effect (350). Wonsik Lee, a post doc in our lab, has shown that our WAP compound series is bactericidal, causing rapid lysis of gram positive bacteria, including *S. aureus*.

We treated the transposon library with these compounds at two different concentrations and performed Tn-Seq on the extracted DNA. Then we put the resulting data through both the data pipelines described above. The upregulation analysis did not reveal any hits for any of the WAP antibiotics at any combinations. This could be because the concentrations at which the transposon library was treated were too low, and there was not enough selective to distinguish a hit from the background noise. On the other hand, we may not identify upregulated genes

because the most robust mechanisms of resistance are due to inactivation of other genes in the genome.

Next, we put the data through the mechanism prediction algorithm. Both concentrations of WAP1 were predicted to be a PG synthesis inhibitor. The lower concentrations of WAP2 and WAP3 were unable to be classified, likely because the concentration used to treat the library was too low. However, the higher concentrations of WAP2 and WAP3 are predicted to bind to Lipid II to inhibit PG synthesis (Table 12). The fact that all of the WAP compounds are predicted to inhibit PG synthesis in one way or another suggests that this could be the mechanism by which this compound series acts. We have validated this prediction using the macromolecular radiolabeling assay (Figure 43). The next steps will be to narrow down the target from the list of many that are part of peptidoglycan biosynthesis. It will also be important to remember that other related pathways can be inhibited, but that the cell still dies due to the inhibition of PG synthesis (ex: daptomycin, bacitracin, and targocil/WTI-11). In either case, these preliminary results suggest that our analysis is working and that we can predict the mechanism of action of antibiotics using Tn-Seq. More experiments are needed to validate this hypothesis.

Table 12. Predicted mechanism of WAP compound series

Compound	Concentration (µg/ml)	Predicted Mechanism
WAP1	0.125	Lipid II Binding
	0.25	Lipid II Binding
WAP2	0.125	Uncategorized
	0.5	PG Synthesis
WAP3	0.125	Uncategorized
	0.25	PG Synthesis
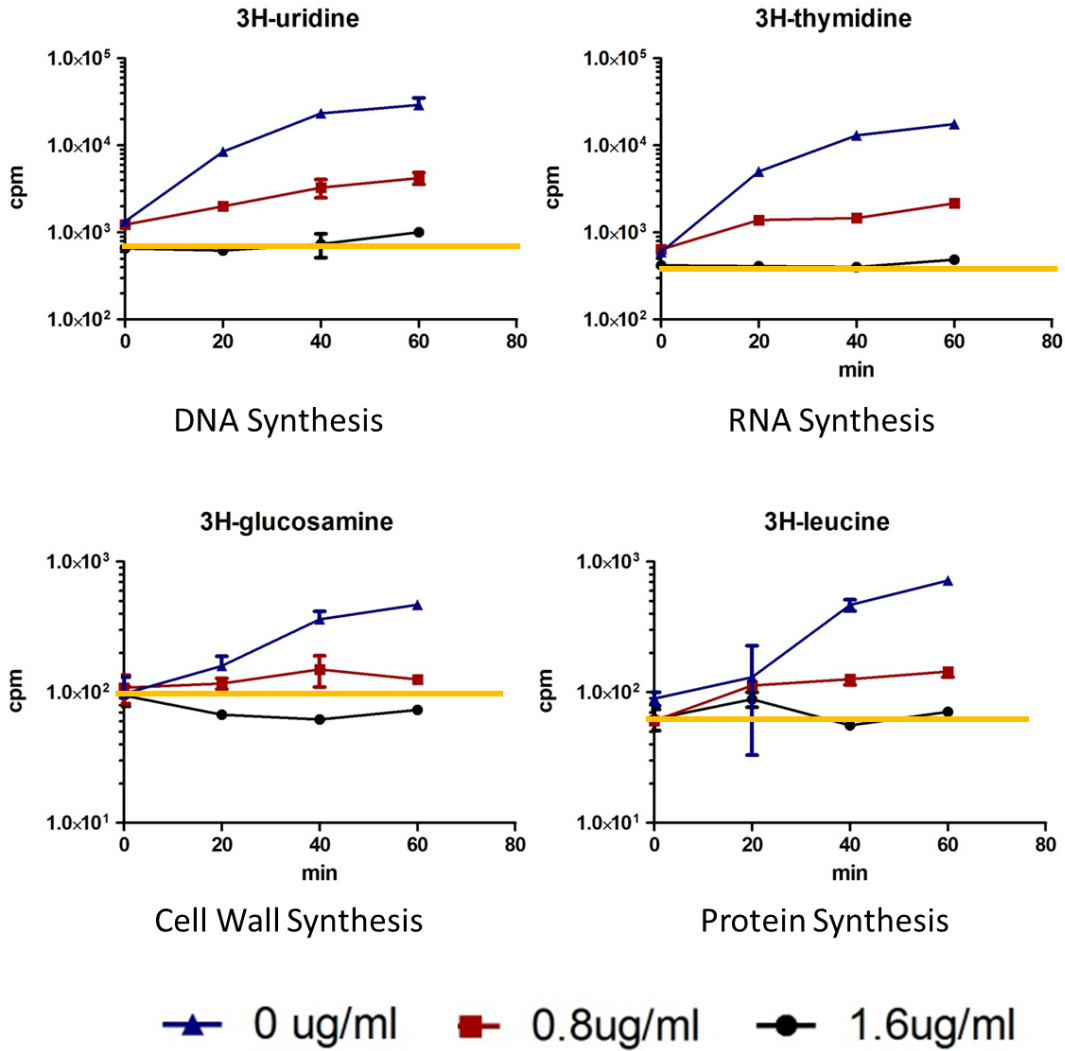
Figure 43. Macromolecular labeling predicts that the WAP compounds kill the cell by inhibiting cell wall synthesis. Macromolecular labeling assays were performed as described (285). Radiolabeled cells were treated with two concentrations of one of the WAP compounds. Though the amount of radiolabel incorporation increased for the radiolabels reporting on DNA synthesis, RNA synthesis, and protein synthesis, the radiolabel reporting on cell wall synthesis decreased at the higher concentration, suggesting that these compounds can inhibit cell wall synthesis. To assist with the comparison, the yellow lines mark the starting level of the radiolabel.

4.6 Discussion

Intrinsic resistance factors of *S. aureus*

In this work, we identified 80 genes that when inactivated can increase or decrease *S. aureus* sensitivity to six different antibiotics. We chose to validate a selection of the hits that were found in more than one antibiotic treatment: *ndh, fmtA,* and *mprF*. Though they all sensitize to different antibiotics, the mechanism of sensitization could be different. *mprF* modifies characteristics of the cell envelope by modifying the charge which can affect how well the antibiotics enter the cell (51). On the other hand, genes like *ndh* are part of global systems that regulate cellular metabolism such as oxidative phosphorylation (299). When these are inactivated, the cell is forced to switch to alternative modes of growth (218) which could sensitize the cell to some antibiotics, or simply slow down growth enough, that the mutant is out-competed by other mutants in the library. Finally, these resistance factors may represent actual synthetic lethal interactions with the target of the antibiotic (189). At sub-MIC antibiotic concentrations, the target of the antibiotic will be partially inhibited. During this partial inhibition, other genes will become essential. These conditionally-essential genes could be redundant genes, genes in parallel pathways, or even genes that interact with the antibiotic's target. Though this analysis does not identify why the resistance factors promote resistance to each antibiotic, we are able to reliably confirm our sequencing results with these experiments.

Global systems as resistance factors

This method also confirms that global systems are very common resistance factors. When components of these systems are inactivated, resistance to antibiotics can increase or decrease. We confirmed that the *agr* system and the *sigB* locus as important resistance factors for oxacillin (and daptomycin for *agr*). The *agr* system is *S. aureus*'s quorum sensing system, but it also regulates genes involved in virulence and antibiotic resistance (306-308). It is

possible that misregulation of some autolysins could lead to the sensitization affect we see with oxacillin and daptomycin. The *sigB* locus consists of *sigB* (the sigma factor)*, rsbW* (the anti-sigma factor)*,* and *rsbV* (the anti-anti-sigma factor) (351, 352). All three genes are sensitive to oxacillin treatment. This locus is involved in regulation of growth phase, the heat shock response, and expression of virulence factors (353) (354). *SigB* has been implicated in resistance to cell-wall targeting antibiotics, and when over-expressed it causes a thickening of the cell wall (296, 355). Likely, it regulates the expression of genes that can modify the cell envelope to defend the cell from oxacillin-induced cell wall stress. Therefore, when *sigB* is inactivated, the cell cannot effectively respond to cell wall stress.

We only identified one global system that can increase resistance when inactivated. Ten genes in the oxidative phosphorylation pathway conferred resistance to gentamicin when knocked out by a transposon insertion. When these genes are interrupted, the cell switches to anaerobic growth, which not only slows down the growth rate of the cell, but also decreases the membrane potential. Without the membrane potential, gentamicin cannot enter the cell, and therefore cannot inhibit its target, the ribosome (294). It seemed remarkable that a set of mutants that slow growth had such a fitness advantage in the presence of an antibiotic that they outcompeted the majority of the other mutants in the library (Table 13). These mutants often cause highly persistent antibiotic-resistant infections with the characteristic small colony variant (SCV) phenotype (218). Our results agree with others that suggest that SCV strains can be induced by gentamicin treatment (356).

Table 13. Increase in percent reads mapping to top hits in oxidative phosphorylation pathway

	% of Total Reads Untreated Control	% of Total Reads Gentamicin Treatment
SAOUHSC_01001	0.00181	3.49398
SAOUHSC_01002	0.00103	0.36541
SAOUHSC_00878	0.00540	23.32021
SAOUHSC_00982	0.00317	9.65116
SAOUHSC_01960	0.00003	0.02395
SAOUHSC_01962	0.00044	0.18194
SAOUHSC_01772	0.00052	1.54628
SAOUHSC_01776	0.00041	0.45655

Role of multi-component sensory systems

MCSs in *S. aureus* mediate responses to many types of conditions including quorum sensing, osmolarity, nutrient availability, and antibiotic treatment(305). Our data suggest that the *graRS/vraFG* MCS is very important for antibiotic resistance. This MCS is known to be important for a variety of cell functions (303, 304), but especially for regulating the charge of the cell envelope through the expression of *dltA* and *mprF.* Only three of the six antibiotics tested (vancomycin, gentamicin, and ciprofloxacin) contain at least one positive charge, but we did not observe a correlation between the number of charges on the antibiotic and the fitness of these strains. Furthermore, in the case of daptomycin, *mprF* inactivation sensitizes the cell, but inactivation of *graR* and *dltA* does not. If the only role of this MCS was only in modifying the charge of the cell envelope in response to positivitely charged antibiotics, we would not expect *mprF* inactivation to cause such a dramatic sensitization. Our results implicate *graRS/vraFG* in having important roles outside of simply modifying the charge of the cell envelope, and it seems

likely that the sensitization observed when this MCS is interrupted is due to a variety of factors including *dltA* and *mprF*.

Machine learning can predict pathways

We used the K-nearest neighbors machine learning algorithm to identify genes that had similar patterns of resistance as the two hypothetical genes we were interested in, *SAOUHSC_01025* and *SAOUHSC_01050.* We hypothesized that this method could help us learn more about the function of hypothetical genes because we previously showed in Chapter 3 that we were able to predict physical interactions by identifying genes with similar resistance and sensitivity patterns across different antibiotics. Using this algorithm, we found components of the *graRS/vraFG* MCS as well as a member of its regulon, *mprF,* suggesting that these genes are important for resistance to cell envelope stresses. There are a few ways that these genes could function that could produce the same resistance pattern as *graRS/vraFG.* If they physically interacted with the *graRS/vraFG* complex, they would have the same pattern as those genes. Also, if it were part of the *graRS/vraFG* regulon, it could have a similar pattern, as we see with *mprF.* The regulon of this system has been well-studied (304), and these genes have not been identified as part of it, so this explanation seems unlikely. Finally, these genes could be part of a parallel and complementary system which, in a semi-redundant manner, protects from cell envelope stress. Though this method cannot predict exactly which pathways these genes are part of, it does help us nominate targets and formulate hypotheses that can be tested with other experiments such as the ones that were done here. Furthermore, more data from different antibiotics with a wider variety of targets should allow us to more accurately predict the function of hypothetical genes.

Tn-Seq and identification of upregulated genes

The analysis for identification of genes upregulated due to transposon insertion did not identify upregulated genes every time. This could be due to the fact that not all the antibiotics

were tested at a high enough concentration where there was not enough selective pressure which makes it difficult to pick out the signal for upregulation from the background level of transposon reads. However, it could also be due to the fact that there is an inactivation mutant that is more fit under antibiotic treatment than the upregulation mutant, so that the upregulation mutant gets outcompeted. It is possible that treating and sequencing the libraries at different antibiotic concentrations may yield better results. Moreover, the cutoffs used here were relatively stringent. It is possible that by decreasing the cutoffs, we may be better able to identify upregulated genes. However, this would also lead to more false positives as we get closer to the background level of transposon reads, and it will be more time intensive to identify true hits. Furthermore, as it stands, every gene we have tested has validated, so we can be confident in our other results. If we were to decrease the cutoffs, it would be vital to validate any candidate genes before proceeding with other experiments to predict the target. Though this method works, the method of plating the transposon library on agar plates does seem to more reliably reveal the target of the antibiotic. It should be possible to plate this transposon library on agar plates containing an antibiotic as was done previously (133, 134). Then the cells could be collected and the location of transposon insertions sequencing using NGS following the protocols described here. This will allow one to have access to the convenience of Tn-Seq (as opposed to sequencing single colonies) without the downsides apparent when growing the cells in liquid culture. We did not choose to perform the experiments this way because we have very little of our unknown compounds. Therefore, if amount of compound is not a limiting factor, this library can still be used in this way to identify upregulated genes and possibly predict the target.

Upregulation of small peptides increases daptomycin resistance

With daptomycin treatment, we identified three small proteins (*SAOUHSC_00969, SAOUHSC_02149,* and *SAOUHSC_02164)* that when upregulated increase resistance to daptomycin. The products these genes code for are very small: 107, 175, and 60 amino acids

long respectively. Of these, only *SAOUHSC_02149* has any predicted domains. According to the NCBI gene website, the N-terminal region has homology to a bacterial Pleckstrin-homology (bPH) domain. Little is known of the function of prokaryotic versions of this gene (357), but eukaryotic proteins with this domain are known for binding phosphatidylinositol and for targeting other proteins to membranes (358, 359). *S. aureus* membranes do not contain phosphatidylinositol, but they do contain phosphatidylglycerol (360). In fact, phosphatidylglycerol is required for daptomycin action (278), and depletion of phosphatidylglycerol results in daptomycin resistance in *B. subtilis* (361). Based on this, the simplest explanation for the reason upregulation of this gene causes resistance is that it is sequestering phosphatidylglycerol away from daptomycin, and without available phosphatidylglycerol, the antibiotic can not enter and disrupt the membrane. Without a predicted domain, it is difficult to hypothesize a mechanism of action for the other two genes. However, in previous experiments, upregulation of other small peptides such as these has been observed in response to cell envelope stress (344). Perhaps they function to stabilize the membrane or recruit other factors to specific parts of the cell to combat the stress induced by daptomycin treatment.

Comparison with other computational methods

We correctly predict the mechanism of action for known antibiotics with a success rate of 72%. This rate is as good or better than other recently-published methods for computationally predicting mechanism of action that have been recently published. In 2015, Wildenhain et al. published research where they used a program called SONAR[G] to predict the mechanism of action of 27 compounds in yeast (290). They report that their algorithm correctly predicted the known targets or an associated target pathway for only 16 of the 27 compounds they tested which translates to a success rate of 59%. Another method published by Johnston et al. in 2016 uses a computational retrobiosynthetic analysis in combination with whole genome sequencing

to predict the mechanism of natural products (289). Not only is this analysis more time-intensive than ours, but they do not publish the rate at which they correctly predict mechanism of action.

Predicting the mechanism for a new antibiotic

Here, we have described two complementary techniques for learning about the mechanism of action for a new antibiotic. Neither technique alone is perfect, but used in unison, they are able to correct for each other. In a best case scenario, these techniques will allow one to predict the target of an antibiotic, but even if it doesn't, it will still lead to new information on potential resistance mechanisms that may occur during further testing or in the clinic which makes the analytical techniques described here very valuable.

As a proof of concept, these approaches work well, but there are many things that could be done to improve the results. For the antibiotics where we were unable to identify upregulated genes, we could try treating the transposon library at a slightly different antibiotic concentration. At a slightly higher concentration, there may be enough selective pressure to decrease the number of genes that we map reads to, which would allow us to more easily identify regions containing the signature of upregulation. In addition, treating at a slightly lower concentration could also help, as perhaps at a concentration close to that of the MIC, upregulated genes may have a higher fitness than some of the inactivation mutants that take over the library at higher concentrations. Unfortunately, as of yet, there is no way to predict what antibiotic concentration will result in the data that best produces upregulated genes. The machine learning approach is similar to other "big data" approaches in that these techniques improve quite a bit as more data is added to the training set. Currently, we have only sequenced one or two antibiotics per target. With more antibiotics per target, we could have a robust training data set that would allow us to more accurately predict the activity of a new antibiotic. Unfortunately, for some targets and even some mechanisms of action, there are not very many antibiotics available.

We are attempting to use this platform to identify the target of a series of related compounds, WAP1-3. Our approach predicts that these compounds inhibit a step in peptidoglycan biosynthesis, and preliminary biochemical experiments support this hypothesis. The next steps include testing whether these drugs bind some precursor in PG biosynthesis or whether they act more like daptomycin, inhibiting PG biosynthesis by binding to and disrupting the membrane (19). Wonsik Lee, a post doc in our lab, is currently performing these experiments.

Conclusions

We have shown that our functional genomics approach to better understanding intrinsic antibiotic resistance can identify known and novel resistance factors. We have validated a number of both the known genes as well as the two hypothetical genes, *SAOUHSC_01025* and *SAOUHSC_01050.* We also propose and utilize a method for predicting the function of hypothetical genes using machine learning to identify known genes with a similar resistance pattern. We have validated this method for these two genes and show that they are sensitive to cell envelope-targeting antibiotics. Any of these validated intrinsic resistance factors are potential targets for the development of potentiating compounds. These hypothetical drugs could be given in combination with currently FDA-approved antibiotics to sensitize *S. aureus* strains. Moreover, this approach can be applied to any antibiotic, and could be useful for not only better understanding the intrinsic resistance factors for that antibiotic, but also for predicting mechanisms of resistance due to gene inactivation. Applying this method to any compound is a robust way to predict intrinsic resistance factors for that compound as well as to learn more about bacterial physiology.

Furthermore, we have also developed two complementary approaches for learning more about antibiotic mechanism of action using Tn-Seq. The advantages of this approach are that with one simple experiment which uses very little compound, it allows anyone to develop

hypotheses about a comopund's mechanism of action that can then be tested with biochemical experiments. This has the potential to speed up and decrease the cost of identifying the target of a new antibiotic. Moreover, we not only learn about the possible targets of the compound, but we learn about possible modes of resistance that could be observed in the clinic due to both inactivation and upregulation of genes. This could allow researchers to begin to address concerns regarding resistance mechanisms early on in the process of developing the antibiotic.

Chapter 5. Summary and Conclusions

Development of new genomics tools for *S. aureus*

Transposon libraries are extremely useful and valuable technologies for better understanding bacterial physiology (89). The methods that existed for making transposon libraries in *S. aureus* were not very reliable, and included a high-temperature plasmid curing step which killed any temperature-sensitive mutant (125). The phage-based high-frequency of transposition platform described here solves these problems (133). We have adapted this system so that we can make a transposon library in any strain of *S. aureus* transducible by Φ11 phage. Furthermore, we modified the standard DNA preparation procedure such that we can now analyze the results of experiments using NGS (132).

The ability to assess and compare the fitness of mutations in every gene in the genome at once gives us a reliable and unbiased method for identifying the most important genes for survival in any condition. To do this, we figured out how to de-multiplex each transposon construct to analyze each one separately. Our analyses allow us to not only identify genes with a statistically-significant change in number of reads which map to them, but we can also identify genes that when upregulated by transposon insertion, confer an increase in fitness. This is especially-interesting because upregulation of genes is a common mechanism of antibiotic resistance. We have validated this platform by identifying the essential genes and comparing them to two other published lists of essential genes (90, 146). Furthermore, we identify temperature-sensitive genes that had previously been annotated as essential (132). Finally, we have also created transposon libraries in other *S. aureus* strains, such as the two community-acquired MRSA strains, MW2 and USA300. This platform has the potential to be very useful for many experiments in the future including but not limited to identification of virulence factors,

factors essential for co-infection with another bacterium, and for comparative genomics studies between different strains of *S. aureus.*

Discovery of new biology

Though the focus of this work was not to answer any particular biological question, in the process of validating the platform and optimizing the analyses, we identified many interesting genes that could begin to explain *S. aureus*'s success as a pathogen. During our temperature experiments, we identified transposon insertions into the menaquinone biosynthetic pathway that were better able to survive high-temperature stress than other mutants in the library. Furthermore, while studying the profiles of genes involved in β-lactam resistance in MRSA strains, we identified a set of genes that appear to be involved in sortase-anchored surface protein secretion. This allowed us to predict a function for a *lytD* that at that point had been completely unstudied (now, its enzymatic activity has been confirmed (197)). We treated transposon libraries with 25 different antibiotics, and for six of them, we have identified the most important intrinsic resistant factors. We were able to identify resistance factors across different classes of antibiotics which could potentially be targets for the development of potentiators to currently-used antibiotics, allowing them to be used even if the bacterium is resistant to the antibiotic alone. Moreover, some of these resistance factors were unstudied genes. We are beginning to characterize some of these genes, *SAOUHSC_01025* and *SAOUHSC_01050*, and a better understanding of their function will lead to a deeper understanding of the actual processes and genes involved in the stress responses to treatment with different antibiotics. Now that this platform for treating and sequencing transposon libraries is up and running in our lab, and we show that it can uncover new biology, we hope that many questions about antibiotic resistance and *S. aureus* physiology will be answered. This could lead to novel strategies and methods for treating antibiotic-resistant infections caused by this dangerous pathogen.

New prediction methods

The success of this system thus far has allowed us to obtain an incredibly large and rich dataset full of potential projects. One major advantage of "big data" is the ability to make predictions based on patterns in a dataset. We devised an approach based on observations we made in Chapter 3, where two genes, *lyrA* and *lytD,* which had very similar resistance and sensitization patterns to three β-lactams, were found to physically interact. In Chapter 4, we have expanded and automated this approach by utilizing machine learning to make predictions. The first method we describe uses six different antibiotics to predict the function of hypothetical genes. With this smaller dataset, we are not able to accurately predict the function of a new gene or its interacting partners, but we can broadly identify which pathways it may be involved in. With a larger dataset, we may be able to more precisely identify a new gene's function. We use a similar approach for predicting antibiotic mechanism of action. We use the resistance and sensitivity fingerprints of the 25 different antibiotics to train the machine learning dataset. This allows us to take a new antibiotic of unknown mechanism and predict which class of antibiotics it is most similar to. Currently, this method classifies antibiotics accurately with an 80% success rate. We expect that with a larger training set consisting of more antibiotics per target, this rate will only improve. Mechanism of action prediction is known to be difficult, and this tool will allow us to begin to nominate targets for an antibiotic that can be validated with follow-up biochemical experiments. Furthermore, antibiotics with novel mechanisms and targets are in high demand. It is possible that this method could identify these novel mechanisms if a new antibiotic is not similar to any antibiotic in our reference dataset.

The future

These approaches are useful for answering not only specific biological questions, but also taking a more systems-level approach to understand bacterial physiology and antibiotic resistance. If we are going to be able to combat dangerous infections, we will need to make use

of novel strategies and new approaches. . Big datasets such as the ones obtained here are incredibly rich, and it would be difficult for one lab can fully make use of them. If this data could be shared with other researchers and other institutions, others viewing it with different eyes may be able to make connections that we never will. I have designed a website where this data can be easily accessed and viewed, and a software developer has assisted me with putting it online. Currently, this is only available for members of the Walker lab, but I have high hopes for a more open-access day in the future when we can make these data public, because only by sharing data and working together will we be able to solve the problem of antibiotic resistance.

References Cited

1.   Trefouel J, Trefouel J, Nitti F, Bovet D. The activity of p-aminophenylsulfamide on experimental streptococcus infections of the mouse and the rabbit. Cr Soc Biol. 1935;120:756-8. PubMed PMID: WOS:000200966300314.

2.   Fleming A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to Their Use in the Isolation of B. Influenzae. Brit J Exp Pathol. 1929;10(3):226-36. PubMed PMID: WOS:000207180600008.

3.   Lewis K. Platforms for antibiotic discovery. Nat Rev Drug Discov. 2013;12(5):371-87. doi: 10.1038/nrd3975. PubMed PMID: WOS:000318350900015.

4.   Clatworthy AE, Pierson E, Hung DT. Targeting virulence: a new paradigm for antimicrobial therapy. Nature chemical biology. 2007;3(9):541-8. doi: 10.1038/nchembio.2007.24. PubMed PMID: 17710100.

5.   McDermott PF, Walker RD, White DG. Antimicrobials: modes of action and mechanisms of resistance. Int J Toxicol. 2003;22(2):135-43. PubMed PMID: 12745995.

6.   Pantosti A, Sanchini A, Monaco M. Mechanisms of antibiotic resistance in Staphylococcus aureus. Future Microbiol. 2007;2(3):323-34. doi: 10.2217/17460913.2.3.323. PubMed PMID: 17661706.

7.   Ubukata K, Itoh-Yamashita N, Konno M. Cloning and expression of the norA gene for fluoroquinolone resistance in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1989;33(9):1535-9. PubMed PMID: 2817852; PMCID: PMC172697.

8.   Cao M, Bernat BA, Wang Z, Armstrong RN, Helmann JD. FosB, a cysteine-dependent fosfomycin resistance protein under the control of sigma(W), an extracytoplasmic-function sigma factor in Bacillus subtilis. Journal of bacteriology. 2001;183(7):2380-3. doi: 10.1128/JB.183.7.2380-2383.2001. PubMed PMID: 11244082; PMCID: PMC95149.

9.   Chambers HF. Methicillin-resistant staphylococci. Clinical microbiology reviews. 1988;1(2):173-86. Epub 1988/04/01. PubMed PMID: 3069195; PMCID: 358041.

10.   Achari A, Somers DO, Champness JN, Bryant PK, Rosemond J, Stammers DK. Crystal structure of the anti-bacterial sulfonamide drug target dihydropteroate synthase. Nat Struct Biol. 1997;4(6):490-7. PubMed PMID: 9187658.

11.   Frere JM. Mechanism of action of beta-lactam antibiotics at the molecular level. Biochem Pharmacol. 1977;26(23):2203-10. PubMed PMID: 337974.

12.   Lando D, Cousin MA, Ojasoo T, Raymond JP. Paromomycin and dihydrostreptomycin binding to Escherichia coli ribosomes. Eur J Biochem. 1976;66(3):597-606. PubMed PMID: 60235.

13.   Ruusala T, Kurland CG. Streptomycin preferentially perturbs ribosomal proofreading. Mol Gen Genet. 1984;198(2):100-4. PubMed PMID: 6394958.

14.   Alvarez-Elcoro S, Enzler MJ. The macrolides: erythromycin, clarithromycin, and azithromycin. Mayo Clin Proc. 1999;74(6):613-34. doi: 10.4065/74.6.613. PubMed PMID: 10377939.

15.   Perkins HR, Nieto M. The chemical basis for the action of the vancomycin group of antibiotics. Annals of the New York Academy of Sciences. 1974;235(0):348-63. Epub 1974/05/10. PubMed PMID: 4369274.

16.   Barna JC, Williams DH. The structure and mode of action of glycopeptide antibiotics of the vancomycin group. Annu Rev Microbiol. 1984;38:339-57. doi: 10.1146/annurev.mi.38.100184.002011. PubMed PMID: 6388496.

17.   Shinabarger D. Mechanism of action of the oxazolidinone antibacterial agents. Expert Opin Investig Drugs. 1999;8(8):1195-202. doi: 10.1517/13543784.8.8.1195. PubMed PMID: 15992144.

18.   Silverman JA, Perlmutter NG, Shapiro HM. Correlation of daptomycin bactericidal activity and membrane depolarization in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2003;47(8):2538-44. PubMed PMID: 12878516; PMCID: PMC166110.

19.   Pogliano J, Pogliano N, Silverman JA. Daptomycin-mediated reorganization of membrane architecture causes mislocalization of essential cell division proteins. Journal of bacteriology. 2012;194(17):4494-504. Epub 2012/06/05. doi: 10.1128/JB.00011-12. PubMed PMID: 22661688; PMCID: 3415520.

20.   Brodersen DE, Clemons WM, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. Cell. 2000;103(7):1143-54. doi: Doi 10.1016/S0092-8674(00)00216-6. PubMed PMID: WOS:000166040600016.

21.   Marconi RT, Lodmell JS, Hill WE. Identification of a Ribosomal-Rna Chloramphenicol Interaction Site within the Peptidyltransferase Center of the 50-S Subunit of the Escherichia-Coli Ribosome. Journal of Biological Chemistry. 1990;265(14):7894-9. PubMed PMID: WOS:A1990DC83500033.

22.   Davies J. Where have All the Antibiotics Gone? Can J Infect Dis Med Microbiol. 2006;17(5):287-90. PubMed PMID: 18382641; PMCID: PMC2095086.

23.   Nathan C. Antibiotics at the crossroads. Nature. 2004;431(7011):899-902. doi: 10.1038/431899a. PubMed PMID: 15496893.

24.   Spellberg B, Guidos R, Gilbert D, Bradley J, Boucher HW, Scheld WM, Bartlett JG, Edwards J, Jr., Infectious Diseases Society of A. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2008;46(2):155-64. doi: 10.1086/524891. PubMed PMID: 18171244.

25.   Hwang TJ, Powers JH, Carpenter D, Kesselheim AS. Accelerating innovation in rapid diagnostics and targeted antibacterials. Nat Biotechnol. 2015;33(6):589-90. doi: 10.1038/nbt.3251. PubMed PMID: 26057972.

26.   Bax R, Green S. Antibiotics: the changing regulatory and pharmaceutical industry paradigm. The Journal of antimicrobial chemotherapy. 2015;70(5):1281-4. doi: 10.1093/jac/dku572. PubMed PMID: 25634991.

27.   Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B, Bartlett J. Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. Clinical Infectious Diseases. 2009;48(1):1-12. doi: 10.1086/595011. PubMed PMID: WOS:000261546800001.

28.   Antibiotic Resistance Threats in the United States 2013. Available from: http://www.cdc.gov/drugresistance/threat-report-2013/pdf/ar-threats-2013-508.pdf.

29.   Kluytmans J, van Belkum A, Verbrugh H. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. Clinical microbiology reviews. 1997;10(3):505-20. PubMed PMID: 9227864; PMCID: PMC172932.

30.   Kuehnert MJ, Kruszon-Moran D, Hill HA, McQuillan G, McAllister SK, Fosheim G, McDougal LK, Chaitram J, Jensen B, Fridkin SK, Killgore G, Tenover FC. Prevalence of Staphylococcus aureus nasal colonization in the United States, 2001-2002. The Journal of infectious diseases. 2006;193(2):172-9. doi: 10.1086/499632. PubMed PMID: 16362880.

31.   Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. Staphylococcus aureus infections: epidemiology, pathophysiology, clinical manifestations, and management. Clinical microbiology reviews. 2015;28(3):603-61. doi: 10.1128/CMR.00134-14. PubMed PMID: 26016486; PMCID: PMC4451395.

32.   Duthie ES, Lorenz LL. Staphylococcal coagulase; mode of action and antigenicity. J Gen Microbiol. 1952;6(1-2):95-107. doi: 10.1099/00221287-6-1-2-95. PubMed PMID: 14927856.

33.   Bae T, Banger AK, Wallace A, Glass EM, Aslund F, Schneewind O, Missiakas DM. Staphylococcus aureus virulence genes identified by bursa aurealis mutagenesis and nematode killing. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(33):12312-7. doi: 10.1073/pnas.0404728101. PubMed PMID: 15304642; PMCID: PMC514475.

34.   Dinges MM, Orwin PM, Schlievert PM. Exotoxins of Staphylococcus aureus. Clinical microbiology reviews. 2000;13(1):16-34, table of contents. PubMed PMID: 10627489; PMCID: PMC88931.

35.   Becker K, Friedrich AW, Lubritz G, Weilert M, Peters G, Von Eiff C. Prevalence of genes encoding pyrogenic toxin superantigens and exfoliative toxins among strains of Staphylococcus aureus isolated from blood and nasal specimens. J Clin Microbiol. 2003;41(4):1434-9. PubMed PMID: 12682126; PMCID: PMC153929.

36.   Lina G, Piemont Y, Godail-Gamot F, Bes M, Peter MO, Gauduchon V, Vandenesch F, Etienne J. Involvement of Panton-Valentine leukocidin-producing Staphylococcus aureus in primary skin infections and pneumonia. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1999;29(5):1128-32. doi: 10.1086/313461. PubMed PMID: 10524952.

37.   Clauditz A, Resch A, Wieland KP, Peschel A, Gotz F. Staphyloxanthin plays a role in the fitness of Staphylococcus aureus and its ability to cope with oxidative stress. Infect Immun. 2006;74(8):4950-3. doi: 10.1128/IAI.00204-06. PubMed PMID: 16861688; PMCID: PMC1539600.

38.   Liu GY, Essex A, Buchanan JT, Datta V, Hoffman HM, Bastian JF, Fierer J, Nizet V. Staphylococcus aureus golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. J Exp Med. 2005;202(2):209-15. doi: 10.1084/jem.20050846. PubMed PMID: 16009720; PMCID: PMC2213009.

39.   International Working Group on the Classification of Staphylococcal Cassette Chromosome E. Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. Antimicrobial agents and chemotherapy. 2009;53(12):4961-7. doi: 10.1128/AAC.00579-09. PubMed PMID: 19721075; PMCID: PMC2786320.

40.   Sievert DM, Ricks P, Edwards JR, Schneider A, Patel J, Srinivasan A, Kallen A, Limbago B, Fridkin S, National Healthcare Safety Network T, Participating NF. Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2009-2010. Infect Control Hosp Epidemiol. 2013;34(1):1-14. doi: 10.1086/668770. PubMed PMID: 23221186.

41.   Kallen AJ, Mu Y, Bulens S, Reingold A, Petit S, Gershman K, Ray SM, Harrison LH, Lynfield R, Dumyati G, Townes JM, Schaffner W, Patel PR, Fridkin SK, Active Bacterial Core surveillance MIotEIP. Health care-associated invasive MRSA infections, 2005-2008. JAMA. 2010;304(6):641-8. doi: 10.1001/jama.2010.1115. PubMed PMID: 20699455.

42.   Esposito S, Noviello S, Leone S. Epidemiology and microbiology of skin and soft tissue infections. Curr Opin Infect Dis. 2016;29(2):109-15. doi: 10.1097/QCO.0000000000000239. PubMed PMID: 26779772.

43.   Miller LG, Eisenberg DF, Liu H, Chang CL, Wang Y, Luthra R, Wallace A, Fang C, Singer J, Suaya JA. Incidence of skin and soft tissue infections in ambulatory and inpatient settings, 2005-2010. BMC Infect Dis. 2015;15:362. doi: 10.1186/s12879-015-1071-0. PubMed PMID: 26293161; PMCID: PMC4546168.

44.   Seybold U, Kourbatova EV, Johnson JG, Halvosa SJ, Wang YF, King MD, Ray SM, Blumberg HM. Emergence of community-associated methicillin-resistant Staphylococcus aureus USA300 genotype as a major cause of health care-associated blood stream infections. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2006;42(5):647-56. doi: 10.1086/499815. PubMed PMID: 16447110.

45.   Kahne D, Leimkuhler C, Lu W, Walsh C. Glycopeptide and lipoglycopeptide antibiotics. Chem Rev. 2005;105(2):425-48. doi: 10.1021/cr030103a. PubMed PMID: 15700951.

46.   Courvalin P. Vancomycin resistance in gram-positive cocci. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2006;42 Suppl 1:S25-34. doi: 10.1086/491711. PubMed PMID: 16323116.

47.   Gould IM. VRSA-doomsday superbug or damp squib? Lancet Infect Dis. 2010;10(12):816-8. doi: 10.1016/S1473-3099(10)70259-0. PubMed PMID: 21109164.

48.	Howden BP, Davies JK, Johnson PD, Stinear TP, Grayson ML. Reduced vancomycin susceptibility in Staphylococcus aureus, including vancomycin-intermediate and heterogeneous vancomycin-intermediate strains: resistance mechanisms, laboratory detection, and clinical implications. Clinical microbiology reviews. 2010;23(1):99-139. doi: 10.1128/CMR.00042-09. PubMed PMID: 20065327; PMCID: PMC2806658.

49.	Todd B. Beyond MRSA: VISA and VRSA: what will ward off these pathogens in health care facilities? Am J Nurs. 2006;106(4):28-30. PubMed PMID: 16575232.

50.	Gardete S, Tomasz A. Mechanisms of vancomycin resistance in Staphylococcus aureus. J Clin Invest. 2014;124(7):2836-40. doi: 10.1172/JCI68834. PubMed PMID: 24983424; PMCID: PMC4071404.

51.	Peschel A, Jack RW, Otto M, Collins LV, Staubitz P, Nicholson G, Kalbacher H, Nieuwenhuizen WF, Jung G, Tarkowski A, van Kessel KP, van Strijp JA. Staphylococcus aureus resistance to human defensins and evasion of neutrophil killing via the novel virulence factor MprF is based on modification of membrane lipids with l-lysine. J Exp Med. 2001;193(9):1067-76. PubMed PMID: 11342591; PMCID: PMC2193429.

52.	Rubio A, Conrad M, Haselbeck RJ, G CK, Brown-Driver V, Finn J, Silverman JA. Regulation of mprF by antisense RNA restores daptomycin susceptibility to daptomycin-resistant isolates of Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2011;55(1):364-7. doi: 10.1128/AAC.00429-10. PubMed PMID: 20974866; PMCID: PMC3019644.

53.	Eliopoulos GM, Thauvin C, Gerson B, Moellering RC, Jr. In vitro activity and mechanism of action of A21978C1, a novel cyclic lipopeptide antibiotic. Antimicrobial agents and chemotherapy. 1985;27(3):357-62. PubMed PMID: 3994349; PMCID: PMC176277.

54.	Mishra NN, Yang SJ, Sawa A, Rubio A, Nast CC, Yeaman MR, Bayer AS. Analysis of cell membrane characteristics of in vitro-selected daptomycin-resistant strains of methicillin-resistant Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2009;53(6):2312-8. doi: 10.1128/AAC.01682-08. PubMed PMID: 19332678; PMCID: PMC2687258.

55.	Mishra NN, Bayer AS, Weidenmaier C, Grau T, Wanner S, Stefani S, Cafiso V, Bertuccio T, Yeaman MR, Nast CC, Yang SJ. Phenotypic and genotypic characterization of daptomycin-resistant methicillin-resistant Staphylococcus aureus strains: relative roles of mprF and dlt operons. PloS one. 2014;9(9):e107426. Epub 2014/09/17. doi: 10.1371/journal.pone.0107426. PubMed PMID: 25226591; PMCID: 4166420.

56.	von Eiff C, Proctor RA, Peters G. Small colony variants of Staphylococci: a link to persistent infections. Berl Munch Tierarztl Wochenschr. 2000;113(9):321-5. PubMed PMID: 11042943.

57.	von Eiff C, Peters G, Becker K. The small colony variant (SCV) concept -- the role of staphylococcal SCVs in persistent infections. Injury. 2006;37 Suppl 2:S26-33. doi: 10.1016/j.injury.2006.04.006. PubMed PMID: 16651068.

58.	Garcia LG, Lemaire S, Kahl BC, Becker K, Proctor RA, Denis O, Tulkens PM, Van Bambeke F. Antibiotic activity against small-colony variants of Staphylococcus aureus: review of in vitro, animal and

clinical data. The Journal of antimicrobial chemotherapy. 2013;68(7):1455-64. doi: 10.1093/jac/dkt072. PubMed PMID: 23485724.

59.   McNamara PJ, Proctor RA. Staphylococcus aureus small colony variants, electron transport and persistent infections. International journal of antimicrobial agents. 2000;14(2):117-22. PubMed PMID: 10720801.

60.   Kriegeskorte A, Grubmuller S, Huber C, Kahl BC, von Eiff C, Proctor RA, Peters G, Eisenreich W, Becker K. Staphylococcus aureus small colony variants show common metabolic features in central metabolism irrespective of the underlying auxotrophism. Front Cell Infect Microbiol. 2014;4:141. doi: 10.3389/fcimb.2014.00141. PubMed PMID: 25374845; PMCID: PMC4204524.

61.   Theuretzbacher U. Recent FDA Antibiotic Approvals: Good News and Bad News: The Center for Disease Dynamics Economics and Policy; 2015 [cited 2016]. Available from: http://cddep.org/blog/posts/recent_fda_antibiotic_approvals_good_news_and_bad_news#sthash.AGlT DhTo.dpbs.

62.   Walsh CT, Wencewicz TA. Prospects for new antibiotics: a molecule-centered perspective. The Journal of antibiotics. 2014;67(1):7-22. doi: 10.1038/ja.2013.49. PubMed PMID: 23756684.

63.   Beadle GW, McClintock B. A Genic Disturbance of Meiosis in Zea Mays. Science. 1928;68(1766):433. doi: 10.1126/science.68.1766.433. PubMed PMID: 17833187.

64.   Mc CB. The origin and behavior of mutable loci in maize. Proceedings of the National Academy of Sciences of the United States of America. 1950;36(6):344-55. PubMed PMID: 15430309; PMCID: PMC1063197.

65.   Reznikoff WS. Transposon Tn5. Annu Rev Genet. 2008;42:269-86. doi: 10.1146/annurev.genet.42.110807.091656. PubMed PMID: 18680433.

66.   Shapiro JA. Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proceedings of the National Academy of Sciences of the United States of America. 1979;76(4):1933-7. PubMed PMID: 287033; PMCID: PMC383507.

67.   Ravindran S. Barbara McClintock and the discovery of jumping genes. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(50):20198-9. doi: 10.1073/pnas.1219372109. PubMed PMID: 23236127; PMCID: PMC3528533.

68.   de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet. 2011;7(12):e1002384. doi: 10.1371/journal.pgen.1002384. PubMed PMID: 22144907; PMCID: PMC3228813.

69.   Alibayov B, Baba-Moussa L, Sina H, Zdenkova K, Demnerova K. Staphylococcus aureus mobile genetic elements. Mol Biol Rep. 2014;41(8):5005-18. doi: 10.1007/s11033-014-3367-3. PubMed PMID: 24728610.

70.   Finnegan DJ. Transposable elements. Curr Opin Genet Dev. 1992;2(6):861-7. PubMed PMID: 1335807.

71.   Choi KH, Kim KJ. Applications of transposon-based gene delivery system in bacteria. J Microbiol Biotechnol. 2009;19(3):217-28. PubMed PMID: 19349746.

72.   Kersulyte D, Velapatino B, Dailide G, Mukhopadhyay AK, Ito Y, Cahuayme L, Parkinson AJ, Gilman RH, Berg DE. Transposable element ISHp608 of Helicobacter pylori: nonrandom geographic distribution, functional organization, and insertion specificity. Journal of bacteriology. 2002;184(4):992-1002. PubMed PMID: 11807059; PMCID: PMC134827.

73.   Hochhut B, Waldor MK. Site-specific integration of the conjugal Vibrio cholerae SXT element into prfC. Molecular microbiology. 1999;32(1):99-110. PubMed PMID: 10216863.

74.   Lamberg A, Nieminen S, Qiao M, Savilahti H. Efficient insertion mutagenesis strategy for bacterial genomes involving electroporation of in vitro-assembled DNA transposition complexes of bacteriophage mu. Appl Environ Microbiol. 2002;68(2):705-12. PubMed PMID: 11823210; PMCID: PMC126711.

75.   Jacobson JW, Medhora MM, Hartl DL. Molecular structure of a somatically unstable transposable element in Drosophila. Proceedings of the National Academy of Sciences of the United States of America. 1986;83(22):8684-8. PubMed PMID: 3022302; PMCID: PMC386995.

76.   Hayes F. Transposon-based strategies for microbial functional genomics and proteomics. Annu Rev Genet. 2003;37:3-29. doi: 10.1146/annurev.genet.37.110801.142807. PubMed PMID: 14616054.

77.   Luft FC. Sleeping Beauty jumps to new heights. J Mol Med (Berl). 2010;88(7):641-3. doi: 10.1007/s00109-010-0626-1. PubMed PMID: 20467721.

78.   Akerley BJ, Rubin EJ, Camilli A, Lampe DJ, Robertson HM, Mekalanos JJ. Systematic identification of essential genes by in vitro mariner mutagenesis. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(15):8927-32. PubMed PMID: 9671781; PMCID: PMC21179.

79.   de Lorenzo V, Herrero M, Jakubzik U, Timmis KN. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. Journal of bacteriology. 1990;172(11):6568-72. PubMed PMID: 2172217; PMCID: PMC526846.

80.   Cvitkovitch DG, Gutierrez JA, Behari J, Youngman PJ, Wetz JE, Crowley PJ, Hillman JD, Brady LJ, Bleiweis AS. Tn917-lac mutagenesis of Streptococcus mutans to identify environmentally regulated genes. FEMS Microbiol Lett. 2000;182(1):149-54. PubMed PMID: 10612747.

81.   Butterfield YS, Marra MA, Asano JK, Chan SY, Guin R, Krzywinski MI, Lee SS, MacDonald KW, Mathewson CA, Olson TE, Pandoh PK, Prabhu AL, Schnerch A, Skalska U, Smailus DE, Stott JM, Tsai MI, Yang GS, Zuyderduyn SD, Schein JE, Jones SJ. An efficient strategy for large-scale high-throughput transposon-mediated sequencing of cDNA clones. Nucleic acids research. 2002;30(11):2460-8. PubMed PMID: 12034834; PMCID: PMC117194.

82.   Manoil C. Tagging exported proteins using Escherichia coli alkaline phosphatase gene fusions. Methods Enzymol. 2000;326:35-47. PubMed PMID: 11036633.

83.   Saenz HL, Dehio C. Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. Current opinion in microbiology. 2005;8(5):612-9. doi: 10.1016/j.mib.2005.08.013. PubMed PMID: 16126452.

84.   Geoffroy MC, Floquet S, Metais A, Nassif X, Pelicic V. Large-scale analysis of the meningococcus genome by gene disruption: resistance to complement-mediated lysis. Genome Res. 2003;13(3):391-8. doi: 10.1101/gr.664303. PubMed PMID: 12618369; PMCID: PMC430250.

85.   Mandin P, Fsihi H, Dussurget O, Vergassola M, Milohanic E, Toledo-Arana A, Lasa I, Johansson J, Cossart P. VirR, a response regulator critical for Listeria monocytogenes virulence. Molecular microbiology. 2005;57(5):1367-80. doi: 10.1111/j.1365-2958.2005.04776.x. PubMed PMID: 16102006.

86.   Sassetti CM, Boyd DH, Rubin EJ. Comprehensive identification of conditionally essential genes in mycobacteria. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(22):12712-7. doi: 10.1073/pnas.231275498. PubMed PMID: 11606763; PMCID: PMC60119.

87.   Badarinarayana V, Estep PW, 3rd, Shendure J, Edwards J, Tavazoie S, Lam F, Church GM. Selection analyses of insertional mutants using subgenic-resolution arrays. Nat Biotechnol. 2001;19(11):1060-5. doi: 10.1038/nbt1101-1060. PubMed PMID: 11689852.

88.   Winterberg KM, Luecke J, Bruegl AS, Reznikoff WS. Phenotypic screening of Escherichia coli K-12 Tn5 insertion libraries, using whole-genome oligonucleotide microarrays. Appl Environ Microbiol. 2005;71(1):451-9. doi: 10.1128/AEM.71.1.451-459.2005. PubMed PMID: 15640221; PMCID: PMC544249.

89.   van Opijnen T, Camilli A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. Nat Rev Microbiol. 2013;11(7):435-42. doi: 10.1038/nrmicro3033. PubMed PMID: 23712350; PMCID: PMC3842022.

90.   Valentino MD, Foulston L, Sadaka A, Kos VN, Villet RA, Santa Maria J, Jr., Lazinski DW, Camilli A, Walker S, Hooper DC, Gilmore MS. Genes contributing to Staphylococcus aureus fitness in abscess- and infection-related ecologies. mBio. 2014;5(5):e01729-14. doi: 10.1128/mBio.01729-14. PubMed PMID: 25182329; PMCID: PMC4173792.

91.   Wang N, Ozer EA, Mandel MJ, Hauser AR. Genome-wide identification of Acinetobacter baumannii genes necessary for persistence in the lung. mBio. 2014;5(3):e01163-14. doi: 10.1128/mBio.01163-14. PubMed PMID: 24895306; PMCID: PMC4049102.

92.   Verhagen LM, de Jonge MI, Burghout P, Schraa K, Spagnuolo L, Mennens S, Eleveld MJ, van der Gaast-de Jongh CE, Zomer A, Hermans PW, Bootsma HJ. Genome-wide identification of genes essential for the survival of Streptococcus pneumoniae in human saliva. PloS one. 2014;9(2):e89541. doi: 10.1371/journal.pone.0089541. PubMed PMID: 24586856; PMCID: PMC3934895.

93.   Palace SG, Proulx MK, Lu S, Baker RE, Goguen JD. Genome-wide mutant fitness profiling identifies nutritional requirements for optimal growth of Yersinia pestis in deep tissue. mBio. 2014;5(4). doi: 10.1128/mBio.01385-14. PubMed PMID: 25139902; PMCID: PMC4147864.

94.   Kieser KJ, Baranowski C, Chao MC, Long JE, Sassetti CM, Waldor MK, Sacchettini JC, Ioerger TR, Rubin EJ. Peptidoglycan synthesis in Mycobacterium tuberculosis is organized into networks with varying drug susceptibility. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(42):13087-92. doi: 10.1073/pnas.1514135112. PubMed PMID: 26438867; PMCID: PMC4620856.

95.   Santa Maria JP, Jr., Sadaka A, Moussa SH, Brown S, Zhang YJ, Rubin EJ, Gilmore MS, Walker S. Compound-gene interaction mapping reveals distinct roles for Staphylococcus aureus teichoic acids. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(34):12510-5. Epub 2014/08/12. doi: 10.1073/pnas.1404099111. PubMed PMID: 25104751; PMCID: 4151746.

96.   Yung MC, Park DM, Overton KW, Blow MJ, Hoover CA, Smit J, Murray SR, Ricci DP, Christen B, Bowman GR, Jiao Y. Transposon Mutagenesis Paired with Deep Sequencing of Caulobacter crescentus under Uranium Stress Reveals Genes Essential for Detoxification and Stress Tolerance. Journal of bacteriology. 2015;197(19):3160-72. doi: 10.1128/JB.00382-15. PubMed PMID: 26195598; PMCID: PMC4560278.

97.   Burghout P, Zomer A, van der Gaast-de Jongh CE, Janssen-Megens EM, Francoijs KJ, Stunnenberg HG, Hermans PW. Streptococcus pneumoniae folate biosynthesis responds to environmental CO2 levels. Journal of bacteriology. 2013;195(7):1573-82. doi: 10.1128/JB.01942-12. PubMed PMID: 23354753; PMCID: PMC3624543.

98.   van Opijnen T, Bodi KL, Camilli A. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nature methods. 2009;6(10):767-72. Epub 2009/09/22. doi: 10.1038/nmeth.1377. PubMed PMID: 19767758; PMCID: 2957483.

99.   Goodman AL, Wu M, Gordon JI. Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. Nature protocols. 2011;6(12):1969-80. Epub 2011/11/19. doi: 10.1038/nprot.2011.417. PubMed PMID: 22094732; PMCID: 3310428.

100.   Gawronski JD, Wong SM, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(38):16422-7. doi: 10.1073/pnas.0906627106. PubMed PMID: 19805314; PMCID: PMC2752563.

101.   Langridge GC, Phan MD, Turner DJ, Perkins TT, Parts L, Haase J, Charles I, Maskell DJ, Peters SE, Dougan G, Wain J, Parkhill J, Turner AK. Simultaneous assay of every Salmonella Typhi gene using one million transposon mutants. Genome Res. 2009;19(12):2308-16. doi: 10.1101/gr.097097.109. PubMed PMID: 19826075; PMCID: PMC2792183.

102.   DeJesus MA, Zhang YJ, Sassetti CM, Rubin EJ, Sacchettini JC, Ioerger TR. Bayesian analysis of gene essentiality based on sequencing of transposon insertion libraries. Bioinformatics. 2013;29(6):695-703. doi: 10.1093/bioinformatics/btt043. PubMed PMID: 23361328; PMCID: PMC3597147.

103.   DeJesus MA, Ioerger TR. A Hidden Markov Model for identifying essential and growth-defect regions in bacterial genomes from transposon insertion sequencing data. BMC Bioinformatics. 2013;14:303. doi: 10.1186/1471-2105-14-303. PubMed PMID: 24103077; PMCID: PMC3854130.

104.   Chao MC, Pritchard JR, Zhang YJ, Rubin EJ, Livny J, Davis BM, Waldor MK. High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. Nucleic acids research. 2013;41(19):9033-48. doi: 10.1093/nar/gkt654. PubMed PMID: 23901011; PMCID: PMC3799429.

105.   Pasquina L, Santa Maria JP, Jr., McKay Wood B, Moussa SH, Matano LM, Santiago M, Martin SE, Lee W, Meredith TC, Walker S. A synthetic lethal approach for compound and target identification in Staphylococcus aureus. Nature chemical biology. 2016;12(1):40-5. doi: 10.1038/nchembio.1967. PubMed PMID: 26619249; PMCID: PMC4684722.

106.   Zomer A, Burghout P, Bootsma HJ, Hermans PW, van Hijum SA. ESSENTIALS: software for rapid analysis of high throughput transposon insertion sequencing data. PloS one. 2012;7(8):e43012. doi: 10.1371/journal.pone.0043012. PubMed PMID: 22900082; PMCID: PMC3416827.

107.   Pritchard JR, Chao MC, Abel S, Davis BM, Baranowski C, Zhang YJ, Rubin EJ, Waldor MK. ARTIST: high-resolution genome-wide assessment of fitness using transposon-insertion sequencing. PLoS Genet. 2014;10(11):e1004782. doi: 10.1371/journal.pgen.1004782. PubMed PMID: 25375795; PMCID: PMC4222735.

108.   Blanchard AM, Leigh JA, Egan SA, Emes RD. Transposon insertion mapping with PIMMS - Pragmatic Insertional Mutation Mapping System. Front Genet. 2015;6:139. doi: 10.3389/fgene.2015.00139. PubMed PMID: 25914720; PMCID: PMC4391243.

109.   Solaimanpour S, Sarmiento F, Mrazek J. Tn-seq explorer: a tool for analysis of high-throughput sequencing data of transposon mutant libraries. PloS one. 2015;10(5):e0126070. doi: 10.1371/journal.pone.0126070. PubMed PMID: 25938432; PMCID: PMC4418687.

110.   DeJesus MA, Ambadipudi C, Baker R, Sassetti C, Ioerger TR. TRANSIT--A Software Tool for Himar1 TnSeq Analysis. PLoS Comput Biol. 2015;11(10):e1004401. doi: 10.1371/journal.pcbi.1004401. PubMed PMID: 26447887; PMCID: PMC4598096.

111.   Berger-Bachi B. Insertional inactivation of staphylococcal methicillin resistance by Tn551. Journal of bacteriology. 1983;154(1):479-87. PubMed PMID: 6300037; PMCID: PMC217482.

112.   Berger-Bachi B, Barberis-Maino L, Strassle A, Kayser FH. FemA, a host-mediated factor essential for methicillin resistance in Staphylococcus aureus: molecular cloning and characterization. Mol Gen Genet. 1989;219(1-2):263-9. PubMed PMID: 2559314.

113.   de Lencastre H, Tomasz A. Reassessment of the number of auxiliary genes essential for expression of high-level methicillin resistance in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1994;38(11):2590-8. PubMed PMID: 7872753; PMCID: PMC188247.

114.   Albus A, Arbeit RD, Lee JC. Virulence of Staphylococcus aureus mutants altered in type 5 capsule production. Infect Immun. 1991;59(3):1008-14. PubMed PMID: 1847696; PMCID: PMC258360.

115.   Gustafson JE, Berger-Bachi B, Strassle A, Wilkinson BJ. Autolysis of methicillin-resistant and -susceptible Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1992;36(3):566-72. PubMed PMID: 1320363; PMCID: PMC190558.

116.   Zscheck KK, Murray BE. Genes involved in the regulation of beta-lactamase production in enterococci and staphylococci. Antimicrobial agents and chemotherapy. 1993;37(9):1966-70. PubMed PMID: 8239614; PMCID: PMC188101.

117.   McDevitt D, Francois P, Vaudaux P, Foster TJ. Molecular characterization of the clumping factor (fibrinogen receptor) of Staphylococcus aureus. Molecular microbiology. 1994;11(2):237-48. PubMed PMID: 8170386.

118.   Mei JM, Nourbakhsh F, Ford CW, Holden DW. Identification of Staphylococcus aureus virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. Molecular microbiology. 1997;26(2):399-407. PubMed PMID: 9383163.

119.   Cheng AG, Kim HK, Burts ML, Krausz T, Schneewind O, Missiakas DM. Genetic requirements for Staphylococcus aureus abscess formation and persistence in host tissues. FASEB J. 2009;23(10):3393-404. doi: 10.1096/fj.09-135467. PubMed PMID: 19525403; PMCID: PMC2747682.

120.   Lan L, Cheng A, Dunman PM, Missiakas D, He C. Golden pigment production and virulence gene expression are affected by metabolisms in Staphylococcus aureus. Journal of bacteriology. 2010;192(12):3068-77. doi: 10.1128/JB.00928-09. PubMed PMID: 20400547; PMCID: PMC2901709.

121.   Cucarella C, Solano C, Valle J, Amorena B, Lasa I, Penades JR. Bap, a Staphylococcus aureus surface protein involved in biofilm formation. Journal of bacteriology. 2001;183(9):2888-96. doi: 10.1128/JB.183.9.2888-2896.2001. PubMed PMID: 11292810; PMCID: PMC99507.

122.   Ochman H, Gerber AS, Hartl DL. Genetic applications of an inverse polymerase chain reaction. Genetics. 1988;120(3):621-3. PubMed PMID: 2852134; PMCID: PMC1203539.

123.   Das S, Noe JC, Paik S, Kitten T. An improved arbitrary primed PCR method for rapid characterization of transposon insertion sites. J Microbiol Methods. 2005;63(1):89-94. doi: 10.1016/j.mimet.2005.02.011. PubMed PMID: 16157212.

124.   Pajunen MI, Pulliainen AT, Finne J, Savilahti H. Generation of transposon insertion mutant libraries for Gram-positive bacteria by electroporation of phage Mu DNA transposition complexes. Microbiology. 2005;151(Pt 4):1209-18. Epub 2005/04/09. doi: 10.1099/mic.0.27807-0. PubMed PMID: 15817788.

125.   Bae T, Glass EM, Schneewind O, Missiakas D. Generating a collection of insertion mutations in the Staphylococcus aureus genome using bursa aurealis. Methods Mol Biol. 2008;416:103-16. Epub 2008/04/09. doi: 10.1007/978-1-59745-321-9_7. PubMed PMID: 18392963.

126.   Han J, He L, Shi W, Xu X, Wang S, Zhang S, Zhang Y. Glycerol uptake is important for L-form formation and persistence in Staphylococcus aureus. PloS one. 2014;9(9):e108325. doi: 10.1371/journal.pone.0108325. PubMed PMID: 25251561; PMCID: PMC4177120.

127.   Grundling A, Missiakas DM, Schneewind O. Staphylococcus aureus mutants with increased lysostaphin resistance. Journal of bacteriology. 2006;188(17):6286-97. doi: 10.1128/JB.00457-06. PubMed PMID: 16923896; PMCID: PMC1595375.

128.   Tu Quoc PH, Genevaux P, Pajunen M, Savilahti H, Georgopoulos C, Schrenzel J, Kelley WL. Isolation and characterization of biofilm formation-defective mutants of Staphylococcus aureus. Infect Immun. 2007;75(3):1079-88. doi: 10.1128/IAI.01143-06. PubMed PMID: 17158901; PMCID: PMC1828571.

129.   Le Breton Y, Mistry P, Valdes KM, Quigley J, Kumar N, Tettelin H, McIver KS. Genome-wide identification of genes required for fitness of group A Streptococcus in human blood. Infect Immun. 2013;81(3):862-75. doi: 10.1128/IAI.00837-12. PubMed PMID: 23297387; PMCID: PMC3584890.

130.   Fey PD, Endres JL, Yajjala VK, Widhelm TJ, Boissy RJ, Bose JL, Bayles KW. A genetic resource for rapid and comprehensive phenotype screening of nonessential Staphylococcus aureus genes. mBio. 2013;4(1):e00537-12. Epub 2013/02/14. doi: 10.1128/mBio.00537-12. PubMed PMID: 23404398; PMCID: 3573662.

131.   Schenk S, Laddaga RA. Improved method for electroporation of Staphylococcus aureus. FEMS Microbiol Lett. 1992;73(1-2):133-8. PubMed PMID: 1521761.

132.   Santiago M, Matano LM, Moussa SH, Gilmore MS, Walker S, Meredith TC. A new platform for ultra-high density Staphylococcus aureus transposon libraries. BMC genomics. 2015;16:252. doi: 10.1186/s12864-015-1361-3. PubMed PMID: 25888466; PMCID: PMC4389836.

133.   Wang H, Claveau D, Vaillancourt JP, Roemer T, Meredith TC. High-frequency transposition for determining antibacterial mode of action. Nature chemical biology. 2011;7(10):720-9. Epub 2011/09/06. doi: 10.1038/nchembio.643. PubMed PMID: 21892185.

134.   Meredith TC, Wang H, Beaulieu P, Grundling A, Roemer T. Harnessing the power of transposon mutagenesis for antibacterial target identification and evaluation. Mobile genetic elements. 2012;2(4):171-8. Epub 2012/10/25. doi: 10.4161/mge.21647. PubMed PMID: 23094235; PMCID: 3469428.

135.   Palmer AC, Kishony R. Opposing effects of target overexpression reveal drug mechanisms. Nature communications. 2014;5:4296. doi: 10.1038/ncomms5296. PubMed PMID: 24980690; PMCID: PMC4408919.

136.   Moellering RC, Jr. MRSA: the first half century. The Journal of antimicrobial chemotherapy. 2012;67(1):4-11. Epub 2011/10/20. doi: 10.1093/jac/dkr437. PubMed PMID: 22010206.

137.   Gould IM, David MZ, Esposito S, Garau J, Lina G, Mazzei T, Peters G. New insights into meticillin-resistant Staphylococcus aureus (MRSA) pathogenesis, treatment and resistance. International journal of antimicrobial agents. 2012;39(2):96-104. Epub 2011/12/27. doi: 10.1016/j.ijantimicag.2011.09.028. PubMed PMID: 22196394.

138.   Stryjewski ME, Corey GR. Methicillin-resistant Staphylococcus aureus: an evolving pathogen. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2014;58 Suppl 1:S10-9. doi: 10.1093/cid/cit613. PubMed PMID: 24343827.

139.   Roemer T, Schneider T, Pinho MG. Auxiliary factors: a chink in the armor of MRSA resistance to beta-lactam antibiotics. Current opinion in microbiology. 2013;16(5):538-48. Epub 2013/07/31. doi: 10.1016/j.mib.2013.06.012. PubMed PMID: 23895826.

140.   Berger-Bachi B. Factors affecting methicillin resistance in Staphylococcus aureus. International journal of antimicrobial agents. 1995;6(1):13-21. Epub 1995/09/01. PubMed PMID: 18611679.

141.   Dordel J, Kim C, Chung M, Pardos de la Gandara M, Holden MT, Parkhill J, de Lencastre H, Bentley SD, Tomasz A. Novel determinants of antibiotic resistance: identification of mutated loci in highly methicillin-resistant subpopulations of methicillin-resistant Staphylococcus aureus. mBio. 2014;5(2):e01000. Epub 2014/04/10. doi: 10.1128/mBio.01000-13. PubMed PMID: 24713324; PMCID: 3993859.

142.   Mwangi MM, Kim C, Chung M, Tsai J, Vijayadamodar G, Benitez M, Jarvie TP, Du L, Tomasz A. Whole-genome sequencing reveals a link between beta-lactam resistance and synthetases of the alarmone (p)ppGpp in Staphylococcus aureus. Microb Drug Resist. 2013;19(3):153-9. Epub 2013/05/11. doi: 10.1089/mdr.2013.0053. PubMed PMID: 23659600; PMCID: 3662374.

143.   Qureshi NK, Yin S, Boyle-Vavra S. The role of the Staphylococcal VraTSR regulatory system on vancomycin resistance and vanA operon expression in vancomycin-resistant Staphylococcus aureus. PloS one. 2014;9(1):e85873. Epub 2014/01/24. doi: 10.1371/journal.pone.0085873. PubMed PMID: 24454941; PMCID: 3893269.

144.   Boyle-Vavra S, Yin S, Jo DS, Montgomery CP, Daum RS. VraT/YvqF is required for methicillin resistance and activation of the VraSR regulon in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2013;57(1):83-95. Epub 2012/10/17. doi: 10.1128/AAC.01651-12. PubMed PMID: 23070169; PMCID: 3535960.

145.   Barquist L, Boinett CJ, Cain AK. Approaches to querying bacterial genomes with transposon-insertion sequencing. RNA Biol. 2013;10(7):1161-9. doi: 10.4161/rna.24765. PubMed PMID: 23635712; PMCID: PMC3849164.

146.   Chaudhuri RR, Allen AG, Owen PJ, Shalom G, Stone K, Harrison M, Burgis TA, Lockyer M, Garcia-Lara J, Foster SJ, Pleasance SJ, Peters SE, Maskell DJ, Charles IG. Comprehensive identification of essential Staphylococcus aureus genes using Transposon-Mediated Differential Hybridisation (TMDH). BMC genomics. 2009;10:291. Epub 2009/07/03. doi: 10.1186/1471-2164-10-291. PubMed PMID: 19570206; PMCID: 2721850.

147.   Christiansen MT, Kaas RS, Chaudhuri RR, Holmes MA, Hasman H, Aarestrup FM. Genome-wide high-throughput screening to investigate essential genes involved in methicillin-resistant Staphylococcus aureus Sequence Type 398 survival. PloS one. 2014;9(2):e89018. Epub 2014/02/25. doi: 10.1371/journal.pone.0089018. PubMed PMID: 24563689; PMCID: 3923074.

148.   Kilby NJ, Snaith MR, Murray JA. Site-specific recombinases: tools for genome engineering. Trends Genet. 1993;9(12):413-21. PubMed PMID: 8122308.

149.	Storici F, Coglievina M, Bruschi CV. A 2-microm DNA-based marker recycling system for multiple gene disruption in the yeast Saccharomyces cerevisiae. Yeast. 1999;15(4):271-83. doi: 10.1002/(SICI)1097-0061(19990315)15:4<271::AID-YEA371>3.0.CO;2-U. PubMed PMID: 10206187.

150.	Bigot Y, Brillet B, Auge-Gouillou C. Conservation of Palindromic and Mirror Motifs within Inverted Terminal Repeats of mariner-like Elements. J Mol Biol. 2005;351(1):108-16. doi: 10.1016/j.jmb.2005.05.006. PubMed PMID: 15946679.

151.	Lis JT, Schleif R. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. Nucleic acids research. 1975;2(3):383-9. Epub 1975/03/01. PubMed PMID: 236548; PMCID: 342844.

152.	Georgiou CD, Papapostolou I, Grintzalis K. Protocol for the quantitative assessment of DNA concentration and damage (fragmentation and nicks). Nature protocols. 2009;4(2):125-31. doi: 10.1038/nprot.2008.222. PubMed PMID: 19180084.

153.	Goecks J, Nekrutenko A, Taylor J, Galaxy T. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 2010;11(8):R86. doi: 10.1186/gb-2010-11-8-r86. PubMed PMID: 20738864; PMCID: PMC2945788.

154.	Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol. 2010;Chapter 19:Unit 19 0 1-21. doi: 10.1002/0471142727.mb1910s89. PubMed PMID: 20069535; PMCID: PMC4264107.

155.	Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 2005;15(10):1451-5. doi: 10.1101/gr.4086505. PubMed PMID: 16169926; PMCID: PMC1240089.

156.	Gillaspy A, Worrell V, Orvis J, Roe B, Dyer D, Iandolo J. The Staphylococcus aureus NCTC8325 genome. ASM Press. 2006.

157.	Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19(9):1639-45. doi: 10.1101/gr.092759.109. PubMed PMID: 19541911; PMCID: PMC2752132.

158.	Kaneda T. Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. Microbiol Rev. 1991;55(2):288-302. PubMed PMID: 1886522; PMCID: PMC372815.

159.	Singh VK, Hattangady DS, Giotis ES, Singh AK, Chamberlain NR, Stuart MK, Wilkinson BJ. Insertional inactivation of branched-chain alpha-keto acid dehydrogenase in Staphylococcus aureus leads to decreased branched-chain membrane fatty acid content and increased susceptibility to certain stresses. Appl Environ Microbiol. 2008;74(19):5882-90. doi: 10.1128/AEM.00882-08. PubMed PMID: 18689519; PMCID: PMC2565972.

160.	Nickel M, Homuth G, Bohnisch C, Mader U, Schweder T. Cold induction of the Bacillus subtilis bkd operon is mediated by increased mRNA stability. Mol Genet Genomics. 2004;272(1):98-107. doi: 10.1007/s00438-004-1038-0. PubMed PMID: 15241682.

161.	Zhu K, Ding X, Julotok M, Wilkinson BJ. Exogenous isoleucine and fatty acid shortening ensure the high content of anteiso-C15:0 fatty acid required for low-temperature growth of Listeria monocytogenes. Appl Environ Microbiol. 2005;71(12):8002-7. doi: 10.1128/AEM.71.12.8002-8007.2005. PubMed PMID: 16332779; PMCID: PMC1317320.

162.	Szurmant H, Nelson K, Kim EJ, Perego M, Hoch JA. YycH regulates the activity of the essential YycFG two-component system in Bacillus subtilis. Journal of bacteriology. 2005;187(15):5419-26. doi: 10.1128/JB.187.15.5419-5426.2005. PubMed PMID: 16030236; PMCID: PMC1196008.

163.	Turck M, Bierbaum G. Purification and activity testing of the full-length YycFGHI proteins of Staphylococcus aureus. PloS one. 2012;7(1):e30403. doi: 10.1371/journal.pone.0030403. PubMed PMID: 22276191; PMCID: PMC3262814.

164.	Szurmant H, Mohan MA, Imus PM, Hoch JA. YycH and YycI interact to regulate the essential YycFG two-component system in Bacillus subtilis. Journal of bacteriology. 2007;189(8):3280-9. Epub 2007/02/20. doi: 10.1128/JB.01936-06. PubMed PMID: 17307850; PMCID: 1855854.

165.	Dubrac S, Boneca IG, Poupel O, Msadek T. New insights into the WalK/WalR (YycG/YycF) essential signal transduction pathway reveal a major role in controlling cell wall metabolism and biofilm formation in Staphylococcus aureus. Journal of bacteriology. 2007;189(22):8257-69. doi: 10.1128/JB.00645-07. PubMed PMID: 17827301; PMCID: PMC2168699.

166.	Ohta T, Saito K, Kuroda M, Honda K, Hirata H, Hayashi H. Molecular cloning of two new heat shock genes related to the hsp70 genes in Staphylococcus aureus. Journal of bacteriology. 1994;176(15):4779-83. PubMed PMID: 8045913; PMCID: PMC196305.

167.	Chastanet A, Fert J, Msadek T. Comparative genomics reveal novel heat shock regulatory mechanisms in Staphylococcus aureus and other Gram-positive bacteria. Molecular microbiology. 2003;47(4):1061-73. PubMed PMID: 12581359.

168.	Chatterjee I, Becker P, Grundmeier M, Bischoff M, Somerville GA, Peters G, Sinha B, Harraghy N, Proctor RA, Herrmann M. Staphylococcus aureus ClpC is required for stress resistance, aconitase activity, growth recovery, and death. Journal of bacteriology. 2005;187(13):4488-96. doi: 10.1128/JB.187.13.4488-4496.2005. PubMed PMID: 15968059; PMCID: PMC1151783.

169.	Derre I, Rapoport G, Msadek T. CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria. Molecular microbiology. 1999;31(1):117-31. PubMed PMID: 9987115.

170.	Wozniak DJ, Tiwari KB, Soufan R, Jayaswal RK. The mcsB gene of the clpC operon is required for stress tolerance and virulence in Staphylococcus aureus. Microbiology. 2012;158(Pt 10):2568-76. doi: 10.1099/mic.0.060749-0. PubMed PMID: 22902728; PMCID: PMC4083623.

171.   Zylicz M, Ang D, Liberek K, Georgopoulos C. Initiation of lambda DNA replication with purified host- and bacteriophage-encoded proteins: the role of the dnaK, dnaJ and grpE heat shock proteins. EMBO J. 1989;8(5):1601-8. PubMed PMID: 2527744; PMCID: PMC400992.

172.   Kozarich JW, Strominger JL. A membrane enzyme from Staphylococcus aureus which catalyzes transpeptidase, carboxypeptidase, and penicillinase activities. The Journal of biological chemistry. 1978;253(4):1272-8. PubMed PMID: 624730.

173.   Pinho MG, de Lencastre H, Tomasz A. Cloning, characterization, and inactivation of the gene pbpC, encoding penicillin-binding protein 3 of Staphylococcus aureus. Journal of bacteriology. 2000;182(4):1074-9. Epub 2000/01/29. PubMed PMID: 10648534; PMCID: 94384.

174.   Qiao Y, Lebar MD, Schirner K, Schaefer K, Tsukamoto H, Kahne D, Walker S. Detection of lipid-linked peptidoglycan precursors by exploiting an unexpected transpeptidase reaction. Journal of the American Chemical Society. 2014;136(42):14678-81. Epub 2014/10/08. doi: 10.1021/ja508147s. PubMed PMID: 25291014; PMCID: 4210121.

175.   Blake KL, O'Neill AJ, Mengin-Lecreulx D, Henderson PJ, Bostock JM, Dunsmore CJ, Simmons KJ, Fishwick CW, Leeds JA, Chopra I. The nature of Staphylococcus aureus MurA and MurZ and approaches for detection of peptidoglycan biosynthesis inhibitors. Molecular microbiology. 2009;72(2):335-43. doi: 10.1111/j.1365-2958.2009.06648.x. PubMed PMID: 19298367.

176.   Kullik I, Jenni R, Berger-Bachi B. Sequence of the putative alanine racemase operon in Staphylococcus aureus: insertional interruption of this operon reduces D-alanine substitution of lipoteichoic acid and autolysis. Gene. 1998;219(1-2):9-17. PubMed PMID: 9756984.

177.   Heidrich C, Templin MF, Ursinus A, Merdanovic M, Berger J, Schwarz H, de Pedro MA, Holtje JV. Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli. Molecular microbiology. 2001;41(1):167-78. PubMed PMID: 11454209.

178.   Muchova K, Chromikova Z, Barak I. Control of Bacillus subtilis cell shape by RodZ. Environ Microbiol. 2013;15(12):3259-71. doi: 10.1111/1462-2920.12200. PubMed PMID: 23879732.

179.   Garner EC, Bernard R, Wang W, Zhuang X, Rudner DZ, Mitchison T. Coupled, circumferential motions of the cell wall synthesis machinery and MreB filaments in B. subtilis. Science. 2011;333(6039):222-5. doi: 10.1126/science.1203285. PubMed PMID: 21636745; PMCID: PMC3235694.

180.   Dominguez-Escobar J, Chastanet A, Crevenna AH, Fromion V, Wedlich-Soldner R, Carballido-Lopez R. Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. Science. 2011;333(6039):225-8. doi: 10.1126/science.1203466. PubMed PMID: 21636744.

181.   Henriques AO, Glaser P, Piggot PJ, Moran CP, Jr. Control of cell shape and elongation by the rodA gene in Bacillus subtilis. Molecular microbiology. 1998;28(2):235-47. PubMed PMID: 9622350.

182.   Steele VR, Bottomley AL, Garcia-Lara J, Kasturiarachchi J, Foster SJ. Multiple essential roles for EzrA in cell division of Staphylococcus aureus. Molecular microbiology. 2011;80(2):542-55. doi: 10.1111/j.1365-2958.2011.07591.x. PubMed PMID: 21401734.

183.   Lithgow JK, Ingham E, Foster SJ. Role of the hprT-ftsH locus in Staphylococcus aureus. Microbiology. 2004;150(Pt 2):373-81. doi: 10.1099/mic.0.26674-0. PubMed PMID: 14766915.

184.   Duman R, Ishikawa S, Celik I, Strahl H, Ogasawara N, Troc P, Lowe J, Hamoen LW. Structural and genetic analyses reveal the protein SepF as a new membrane anchor for the Z ring. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(48):E4601-10. doi: 10.1073/pnas.1313978110. PubMed PMID: 24218584; PMCID: PMC3845145.

185.   Claessen D, Emmins R, Hamoen LW, Daniel RA, Errington J, Edwards DH. Control of the cell elongation-division cycle by shuttling of PBP1 protein in Bacillus subtilis. Molecular microbiology. 2008;68(4):1029-46. doi: 10.1111/j.1365-2958.2008.06210.x. PubMed PMID: 18363795.

186.   Liu G, Draper GC, Donachie WD. FtsK is a bifunctional protein involved in cell division and chromosome localization in Escherichia coli. Molecular microbiology. 1998;29(3):893-903. PubMed PMID: 9723927.

187.   Chan YG, Frankel MB, Dengler V, Schneewind O, Missiakas D. Staphylococcus aureus mutants lacking the LytR-CpsA-Psr family of enzymes release cell wall teichoic acids into the extracellular medium. Journal of bacteriology. 2013;195(20):4650-9. doi: 10.1128/JB.00544-13. PubMed PMID: 23935043; PMCID: PMC3807444.

188.   Atilano ML, Pereira PM, Yates J, Reed P, Veiga H, Pinho MG, Filipe SR. Teichoic acids are temporal and spatial regulators of peptidoglycan cross-linking in Staphylococcus aureus. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(44):18991-6. doi: 10.1073/pnas.1004304107. PubMed PMID: 20944066; PMCID: PMC2973906.

189.   Campbell J, Singh AK, Santa Maria JP, Jr., Kim Y, Brown S, Swoboda JG, Mylonakis E, Wilkinson BJ, Walker S. Synthetic lethal compound combinations reveal a fundamental connection between wall teichoic acid and peptidoglycan biosyntheses in Staphylococcus aureus. ACS chemical biology. 2011;6(1):106-16. Epub 2010/10/22. doi: 10.1021/cb100269f. PubMed PMID: 20961110; PMCID: 3025082.

190.   Oku Y, Kurokawa K, Ichihashi N, Sekimizu K. Characterization of the Staphylococcus aureus mprF gene, involved in lysinylation of phosphatidylglycerol. Microbiology. 2004;150(Pt 1):45-51. doi: 10.1099/mic.0.26706-0. PubMed PMID: 14702396.

191.   Caspi R, Altman T, Dale JM, Dreher K, Fulcher CA, Gilham F, Kaipa P, Karthikeyan AS, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Paley S, Popescu L, Pujar A, Shearer AG, Zhang P, Karp PD. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic acids research. 2010;38(Database issue):D473-9. Epub 2009/10/24. doi: 10.1093/nar/gkp875. PubMed PMID: 19850718; PMCID: 2808959.

192.   Bentley R, Meganathan R. Biosynthesis of vitamin K (menaquinone) in bacteria. Microbiol Rev. 1982;46(3):241-80. PubMed PMID: 6127606; PMCID: PMC281544.

193.   Nowicka B, Kruk J. Occurrence, biosynthesis and function of isoprenoid quinones. Biochimica et biophysica acta. 2010;1797(9):1587-605. doi: 10.1016/j.bbabio.2010.06.007. PubMed PMID: 20599680.

194.   Hansson M, Hederstedt L. Cloning and characterization of the Bacillus subtilis hemEHY gene cluster, which encodes protoheme IX biosynthetic enzymes. Journal of bacteriology. 1992;174(24):8081-93. PubMed PMID: 1459957; PMCID: PMC207547.

195.   Baumert N, von Eiff C, Schaaff F, Peters G, Proctor RA, Sahl HG. Physiology and antibiotic susceptibility of Staphylococcus aureus small colony variants. Microb Drug Resist. 2002;8(4):253-60. doi: 10.1089/10766290260469507. PubMed PMID: 12523621.

196.   Margot P, Mauel C, Karamata D. The gene of the N-acetylglucosaminidase, a Bacillus subtilis 168 cell wall hydrolase not involved in vegetative cell autolysis. Molecular microbiology. 1994;12(4):535-45. PubMed PMID: 7934877.

197.   Chan YG, Frankel MB, Missiakas D, Schneewind O. SagB Glucosaminidase Is a Determinant of Staphylococcus aureus Glycan Chain Length, Antibiotic Susceptibility, and Protein Secretion. Journal of bacteriology. 2016;198(7):1123-36. doi: 10.1128/JB.00983-15. PubMed PMID: 26811319.

198.   Bae T, Baba T, Hiramatsu K, Schneewind O. Prophages of Staphylococcus aureus Newman and their contribution to virulence. Molecular microbiology. 2006;62(4):1035-47. doi: 10.1111/j.1365-2958.2006.05441.x. PubMed PMID: 17078814.

199.   Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ, Dunman PM. Characterization of the Staphylococcus aureus heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. Journal of bacteriology. 2006;188(19):6739-56. doi: 10.1128/JB.00609-06. PubMed PMID: 16980476; PMCID: PMC1595530.

200.   Costa CS, Anton DN. Round-cell mutants of Salmonella typhimurium produced by transposition mutagenesis: lethality of rodA and mre mutations. Mol Gen Genet. 1993;236(2-3):387-94. PubMed PMID: 8382342.

201.   Sham LT, Tsui HC, Land AD, Barendt SM, Winkler ME. Recent advances in pneumococcal peptidoglycan biosynthesis suggest new vaccine and antimicrobial targets. Current opinion in microbiology. 2012;15(2):194-203. doi: 10.1016/j.mib.2011.12.013. PubMed PMID: 22280885; PMCID: PMC3322672.

202.   Zapun A, Vernet T, Pinho MG. The different shapes of cocci. FEMS Microbiol Rev. 2008;32(2):345-60. doi: 10.1111/j.1574-6976.2007.00098.x. PubMed PMID: 18266741.

203.   Pinho MG, Kjos M, Veening JW. How to get (a)round: mechanisms controlling growth and division of coccoid bacteria. Nat Rev Microbiol. 2013;11(9):601-14. doi: 10.1038/nrmicro3088. PubMed PMID: 23949602.

204.   Pinho MG, Errington J. Dispersed mode of Staphylococcus aureus cell wall synthesis in the absence of the division machinery. Molecular microbiology. 2003;50(3):871-81. PubMed PMID: 14617148.

205.   Margolin W. Sculpting the bacterial cell. Curr Biol. 2009;19(17):R812-22. doi: 10.1016/j.cub.2009.06.033. PubMed PMID: 19906583; PMCID: PMC4080913.

206.   Nishi H, Komatsuzawa H, Fujiwara T, McCallum N, Sugai M. Reduced content of lysyl-phosphatidylglycerol in the cytoplasmic membrane affects susceptibility to moenomycin, as well as vancomycin, gentamicin, and antimicrobial peptides, in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2004;48(12):4800-7. doi: 10.1128/AAC.48.12.4800-4807.2004. PubMed PMID: 15561859; PMCID: PMC529239.

207.   Grundling A, Schneewind O. Cross-linked peptidoglycan mediates lysostaphin binding to the cell wall envelope of Staphylococcus aureus. Journal of bacteriology. 2006;188(7):2463-72. doi: 10.1128/JB.188.7.2463-2472.2006. PubMed PMID: 16547033; PMCID: PMC1428428.

208.   Schindler CA, Schuhardt VT. Lysostaphin: A New Bacteriolytic Agent for the Staphylococcus. Proceedings of the National Academy of Sciences of the United States of America. 1964;51:414-21. PubMed PMID: 14171453; PMCID: PMC300087.

209.   Frankel MB, Wojcik BM, DeDent AC, Missiakas DM, Schneewind O. ABI domain-containing proteins contribute to surface protein display and cell division in Staphylococcus aureus. Molecular microbiology. 2010;78(1):238-52. doi: 10.1111/j.1365-2958.2010.07334.x. PubMed PMID: 20923422; PMCID: PMC3538852.

210.   Pei J, Mitchell DA, Dixon JE, Grishin NV. Expansion of type II CAAX proteases reveals evolutionary origin of gamma-secretase subunit APH-1. J Mol Biol. 2011;410(1):18-26. doi: 10.1016/j.jmb.2011.04.066. PubMed PMID: 21570408; PMCID: PMC3155266.

211.   Dean MA, Olsen RJ, Long SW, Rosato AE, Musser JM. Identification of point mutations in clinical Staphylococcus aureus strains that produce small-colony variants auxotrophic for menadione. Infect Immun. 2014;82(4):1600-5. doi: 10.1128/IAI.01487-13. PubMed PMID: 24452687; PMCID: PMC3993378.

212.   Biswas L, Biswas R, Schlag M, Bertram R, Gotz F. Small-colony variant selection as a survival strategy for Staphylococcus aureus in the presence of Pseudomonas aeruginosa. Appl Environ Microbiol. 2009;75(21):6910-2. doi: 10.1128/AEM.01211-09. PubMed PMID: 19717621; PMCID: PMC2772425.

213.   Mayfield JA, Hammer ND, Kurker RC, Chen TK, Ojha S, Skaar EP, DuBois JL. The chlorite dismutase (HemQ) from Staphylococcus aureus has a redox-sensitive heme and is associated with the small colony variant phenotype. The Journal of biological chemistry. 2013;288(32):23488-504. doi: 10.1074/jbc.M112.442335. PubMed PMID: 23737523.

214.   Onyango LA, Hugh Dunstan R, Roberts TK, Macdonald MM, Gottfries J. Phenotypic variants of staphylococci and their underlying population distributions following exposure to stress. PloS one. 2013;8(10):e77614. doi: 10.1371/journal.pone.0077614. PubMed PMID: 24204894; PMCID: PMC3799968.

215.   von Eiff C, McNamara P, Becker K, Bates D, Lei XH, Ziman M, Bochner BR, Peters G, Proctor RA. Phenotype microarray profiling of Staphylococcus aureus menD and hemB mutants with the small-colony-variant phenotype. Journal of bacteriology. 2006;188(2):687-93. doi: 10.1128/JB.188.2.687-693.2006. PubMed PMID: 16385058; PMCID: PMC1347289.

216.   von Eiff C, Heilmann C, Proctor RA, Woltz C, Peters G, Gotz F. A site-directed Staphylococcus aureus hemB mutant is a small-colony variant which persists intracellularly. Journal of bacteriology. 1997;179(15):4706-12. PubMed PMID: 9244256; PMCID: PMC179315.

217.   Jensen J. Biosynthesis of hematin compounds in a hemin requiring strain of Micrococcus pyogenes var. aureus. I. The significance of coenzyme A for the terminal synthesis of catalase. Journal of bacteriology. 1957;73(3):324-33. PubMed PMID: 13416192; PMCID: PMC289801.

218.   Proctor RA, von Eiff C, Kahl BC, Becker K, McNamara P, Herrmann M, Peters G. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. Nat Rev Microbiol. 2006;4(4):295-305. doi: 10.1038/nrmicro1384. PubMed PMID: 16541137.

219.   Proctor RA, Balwit JM, Vesga O. Variant subpopulations of Staphylococcus aureus as cause of persistent and recurrent infections. Infect Agents Dis. 1994;3(6):302-12. PubMed PMID: 7889317.

220.   Proctor RA, van Langevelde P, Kristjansson M, Maslow JN, Arbeit RD. Persistent and relapsing infections associated with small-colony variants of Staphylococcus aureus. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1995;20(1):95-102. PubMed PMID: 7727677.

221.   Bates DM, von Eiff C, McNamara PJ, Peters G, Yeaman MR, Bayer AS, Proctor RA. Staphylococcus aureus menD and hemB mutants are as infective as the parent strains, but the menadione biosynthetic mutant persists within the kidney. The Journal of infectious diseases. 2003;187(10):1654-61. doi: 10.1086/374642. PubMed PMID: 12721946.

222.   Clements MO, Watson SP, Poole RK, Foster SJ. CtaA of Staphylococcus aureus is required for starvation survival, recovery, and cytochrome biosynthesis. Journal of bacteriology. 1999;181(2):501-7. PubMed PMID: 9882664; PMCID: PMC93404.

223.   Kohler C, von Eiff C, Peters G, Proctor RA, Hecker M, Engelmann S. Physiological characterization of a heme-deficient mutant of Staphylococcus aureus by a proteomic approach. Journal of bacteriology. 2003;185(23):6928-37. PubMed PMID: 14617657; PMCID: PMC262702.

224.   Wolter DJ, Emerson JC, McNamara S, Buccat AM, Qin X, Cochrane E, Houston LS, Rogers GB, Marsh P, Prehar K, Pope CE, Blackledge M, Deziel E, Bruce KD, Ramsey BW, Gibson RL, Burns JL, Hoffman LR. Staphylococcus aureus small-colony variants are independently associated with worse lung disease in children with cystic fibrosis. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;57(3):384-91. doi: 10.1093/cid/cit270. PubMed PMID: 23625938; PMCID: PMC3888146.

225.   Tan NC, Cooksley CM, Roscioli E, Drilling AJ, Douglas R, Wormald PJ, Vreugde S. Small-colony variants and phenotype switching of intracellular Staphylococcus aureus in chronic rhinosinusitis. Allergy. 2014;69(10):1364-71. doi: 10.1111/all.12457. PubMed PMID: 24922342.

226.   Lechner S, Lewis K, Bertram R. Staphylococcus aureus persisters tolerant to bactericidal antibiotics. J Mol Microbiol Biotechnol. 2012;22(4):235-44. doi: 10.1159/000342449

000342449. PubMed PMID: 22986269; PMCID: PMC3518770.

227. Chuard C, Vaudaux PE, Proctor RA, Lew DP. Decreased susceptibility to antibiotic killing of a stable small colony variant of Staphylococcus aureus in fluid phase and on fibronectin-coated surfaces. The Journal of antimicrobial chemotherapy. 1997;39(5):603-8. PubMed PMID: 9184359.

228. Koch G, Yepes A, Forstner KU, Wermser C, Stengel ST, Modamio J, Ohlsen K, Foster KR, Lopez D. Evolution of resistance to a last-resort antibiotic in Staphylococcus aureus via bacterial competition. Cell. 2014;158(5):1060-71. doi: 10.1016/j.cell.2014.06.046. PubMed PMID: 25171407; PMCID: PMC4163622.

229. Utsui Y, Yokota T. Role of an altered penicillin-binding protein in methicillin- and cephem-resistant Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1985;28(3):397-403. Epub 1985/09/01. PubMed PMID: 3878127; PMCID: 180261.

230. Hartman BJ, Tomasz A. Low-affinity penicillin-binding protein associated with beta-lactam resistance in Staphylococcus aureus. Journal of bacteriology. 1984;158(2):513-6. PubMed PMID: 6563036; PMCID: PMC215458.

231. Leski TA, Tomasz A. Role of penicillin-binding protein 2 (PBP2) in the antibiotic susceptibility and cell wall cross-linking of Staphylococcus aureus: evidence for the cooperative functioning of PBP2, PBP4, and PBP2A. Journal of bacteriology. 2005;187(5):1815-24. Epub 2005/02/18. doi: 10.1128/JB.187.5.1815-1824.2005. PubMed PMID: 15716453; PMCID: 1064008.

232. Zapun A, Contreras-Martel C, Vernet T. Penicillin-binding proteins and beta-lactam resistance. FEMS Microbiol Rev. 2008;32(2):361-85. doi: 10.1111/j.1574-6976.2007.00095.x. PubMed PMID: 18248419.

233. Georgopapadakou NH, Smith SA, Bonner DP. Penicillin-binding proteins in a Staphylococcus aureus strain resistant to specific beta-lactam antibiotics. Antimicrobial agents and chemotherapy. 1982;22(1):172-5. Epub 1982/07/01. PubMed PMID: 7125630; PMCID: 183698.

234. Pinho MG, de Lencastre H, Tomasz A. An acquired and a native penicillin-binding protein cooperate in building the cell wall of drug-resistant staphylococci. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10886-91. doi: 10.1073/pnas.191260798. PubMed PMID: 11517340; PMCID: PMC58569.

235. Pinho MG, Ludovice AM, Wu S, De Lencastre H. Massive reduction in methicillin resistance by transposon inactivation of the normal PBP2 in a methicillin-resistant strain of Staphylococcus aureus. Microb Drug Resist. 1997;3(4):409-13. PubMed PMID: 9442495.

236. Pinho MG, Filipe SR, de Lencastre H, Tomasz A. Complementation of the essential peptidoglycan transpeptidase function of penicillin-binding protein 2 (PBP2) by the drug resistance protein PBP2A in Staphylococcus aureus. Journal of bacteriology. 2001;183(22):6525-31. doi: 10.1128/JB.183.22.6525-6531.2001. PubMed PMID: 11673420; PMCID: PMC95481.

237. Pereira SF, Henriques AO, Pinho MG, de Lencastre H, Tomasz A. Role of PBP1 in cell division of Staphylococcus aureus. Journal of bacteriology. 2007;189(9):3525-31. doi: 10.1128/JB.00044-07. PubMed PMID: 17307860; PMCID: PMC1855886.

238.   Pereira SF, Henriques AO, Pinho MG, de Lencastre H, Tomasz A. Evidence for a dual role of PBP1 in the cell division and cell separation of Staphylococcus aureus. Molecular microbiology. 2009;72(4):895-904. doi: 10.1111/j.1365-2958.2009.06687.x. PubMed PMID: 19400776; PMCID: PMC2771448.

239.   Berti AD, Theisen E, Sauer JD, Nonejuie P, Olson J, Pogliano J, Sakoulas G, Nizet V, Proctor RA, Rose WE. Penicillin Binding Protein 1 Is Important in the Compensatory Response of Staphylococcus aureus to Daptomycin-Induced Membrane Damage and Is a Potential Target for beta-Lactam-Daptomycin Synergy. Antimicrobial agents and chemotherapy. 2015;60(1):451-8. doi: 10.1128/AAC.02071-15. PubMed PMID: 26525797; PMCID: PMC4704217.

240.   White CL, Kitich A, Gober JW. Positioning cell wall synthetic complexes by the bacterial morphogenetic proteins MreB and MreD. Molecular microbiology. 2010;76(3):616-33. doi: 10.1111/j.1365-2958.2010.07108.x. PubMed PMID: 20233306.

241.   Shiomi D, Toyoda A, Aizu T, Ejima F, Fujiyama A, Shini T, Kohara Y, Niki H. Mutations in cell elongation genes mreB, mrdA and mrdB suppress the shape defect of RodZ-deficient cells. Molecular microbiology. 2013;87(5):1029-44. doi: 10.1111/mmi.12148. PubMed PMID: 23301723; PMCID: PMC3599482.

242.   Yoshida H, Kawai F, Obayashi E, Akashi S, Roper DI, Tame JR, Park SY. Crystal structures of penicillin-binding protein 3 (PBP3) from methicillin-resistant Staphylococcus aureus in the apo and cefotaxime-bound forms. J Mol Biol. 2012;423(3):351-64. doi: 10.1016/j.jmb.2012.07.012. PubMed PMID: 22846910.

243.   Loskill P, Pereira PM, Jung P, Bischoff M, Herrmann M, Pinho MG, Jacobs K. Reduction of the peptidoglycan crosslinking causes a decrease in stiffness of the Staphylococcus aureus cell envelope. Biophysical journal. 2014;107(5):1082-9. Epub 2014/09/05. doi: 10.1016/j.bpj.2014.07.029. PubMed PMID: 25185544; PMCID: 4156677.

244.   Wyke AW, Ward JB, Hayes MV, Curtis NA. A role in vivo for penicillin-binding protein-4 of Staphylococcus aureus. Eur J Biochem. 1981;119(2):389-93. PubMed PMID: 7308191.

245.   Henze UU, Berger-Bachi B. Penicillin-binding protein 4 overproduction increases beta-lactam resistance in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1996;40(9):2121-5. PubMed PMID: 8878592; PMCID: PMC163484.

246.   Finan JE, Archer GL, Pucci MJ, Climo MW. Role of penicillin-binding protein 4 in expression of vancomycin resistance among clinical isolates of oxacillin-resistant Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2001;45(11):3070-5. doi: 10.1128/AAC.45.11.3070-3075.2001. PubMed PMID: 11600358; PMCID: PMC90784.

247.   Memmi G, Filipe SR, Pinho MG, Fu Z, Cheung A. Staphylococcus aureus PBP4 is essential for beta-lactam resistance in community-acquired methicillin-resistant strains. Antimicrobial agents and chemotherapy. 2008;52(11):3955-66. Epub 2008/08/30. doi: 10.1128/AAC.00049-08. PubMed PMID: 18725435; PMCID: 2573147.

248.   Sieradzki K, Pinho MG, Tomasz A. Inactivated pbp4 in highly glycopeptide-resistant laboratory mutants of Staphylococcus aureus. The Journal of biological chemistry. 1999;274(27):18942-6. PubMed PMID: 10383392.

249.   Chambers HF. Methicillin resistance in staphylococci: molecular and biochemical basis and clinical implications. Clinical microbiology reviews. 1997;10(4):781-91. PubMed PMID: 9336672; PMCID: PMC172944.

250.   Huang H, Flynn NM, King JH, Monchaud C, Morita M, Cohen SH. Comparisons of community-associated methicillin-resistant Staphylococcus aureus (MRSA) and hospital-associated MSRA infections in Sacramento, California. J Clin Microbiol. 2006;44(7):2423-7. doi: 10.1128/JCM.00254-06. PubMed PMID: 16825359; PMCID: PMC1489486.

251.   Moran GJ, Krishnadasan A, Gorwitz RJ, Fosheim GE, McDougal LK, Carey RB, Talan DA, Group EMINS. Methicillin-resistant S. aureus infections among patients in the emergency department. N Engl J Med. 2006;355(7):666-74. doi: 10.1056/NEJMoa055356. PubMed PMID: 16914702.

252.   Mishaan AMA, Mason EO, Martinez-Aguilar G, Hammerman W, Propst JJ, Lupski JR, Stankiewicz P, Kaplan SL, Hulten K. Emergence of a predominant clone of community-acquired Staphylococcus aureus among children in Houston, Texas. Pediatr Infect Dis J. 2005;24(3):201-6. doi: 10.1097/01.inf.0000151107.29132.70. PubMed PMID: WOS:000227567700002.

253.   Christianson S, Golding GR, Campbell J, Mulvey MR. Comparative genomics of Canadian epidemic lineages of methicillin-resistant Staphylococcus aureus. Journal of Clinical Microbiology. 2007;45(6):1904-11. doi: 10.1128/Jcm.02500-06. PubMed PMID: WOS:000247286500036.

254.   Highlander SK, Hulten KG, Qin X, Jiang H, Yerrapragada S, Mason EO, Shang Y, Williams TM, Fortunov RM, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox GE, Cardenas AM, Muzny DM, Hemphill L, Ding Y, Dugan S, Blyth PR, Buhay CJ, Dinh HH, Hawes AC, Holder M, Kovar CL, Lee SL, Liu W, Nazareth LV, Wang Q, Zhou J, Kaplan SL, Weinstock GM. Subtle genetic changes enhance virulence of methicillin resistant and sensitive Staphylococcus aureus. Bmc Microbiol. 2007;7. doi: Artn 99

10.1186/1471-2180-7-99. PubMed PMID: WOS:000253111300001.

255.   Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. Genome and virulence determinants of high virulence community-acquired MRSA. Lancet. 2002;359(9320):1819-27. PubMed PMID: 12044378.

256.   Kato F, Sugai M. A simple method of markerless gene deletion in Staphylococcus aureus. J Microbiol Methods. 2011;87(1):76-81. doi: 10.1016/j.mimet.2011.07.010. PubMed PMID: 21801759.

257.   Tenover FC, Goering RV. Methicillin-resistant Staphylococcus aureus strain USA300: origin and epidemiology. The Journal of antimicrobial chemotherapy. 2009;64(3):441-6. doi: 10.1093/jac/dkp241. PubMed PMID: 19608582.

258.   Roberts GA, Houston PJ, White JH, Chen K, Stephanou AS, Cooper LP, Dryden DT, Lindsay JA. Impact of target site distribution for Type I restriction enzymes on the evolution of methicillin-resistant

Staphylococcus aureus (MRSA) populations. Nucleic acids research. 2013;41(15):7472-84. doi: 10.1093/nar/gkt535. PubMed PMID: 23771140; PMCID: PMC3753647.

259.   Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev. 2000;64(2):412-34. PubMed PMID: 10839821; PMCID: PMC98998.

260.   Iordanescu S, Surdeanu M. Two restriction and modification systems in Staphylococcus aureus NCTC8325. J Gen Microbiol. 1976;96(2):277-81. doi: 10.1099/00221287-96-2-277. PubMed PMID: 136497.

261.   Stobberingh EE, Winkler KC. Restriction-deficient mutants of Staphylococcus aureus. J Gen Microbiol. 1977;99(2):359-67. doi: 10.1099/00221287-99-2-359. PubMed PMID: 141491.

262.   Engman J, Rogstam A, Frees D, Ingmer H, von Wachenfeldt C. The YjbH adaptor protein enhances proteolysis of the transcriptional regulator Spx in Staphylococcus aureus. Journal of bacteriology. 2012;194(5):1186-94. doi: 10.1128/JB.06414-11. PubMed PMID: 22194450; PMCID: PMC3294810.

263.   Gohring N, Fedtke I, Xia G, Jorge AM, Pinho MG, Bertsche U, Peschel A. New role of the disulfide stress effector YjbH in beta-lactam susceptibility of Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2011;55(12):5452-8. doi: 10.1128/AAC.00286-11. PubMed PMID: 21947404; PMCID: PMC3232775.

264.   Hancock IC, Wiseman G, Baddiley J. Biosynthesis of the unit that links teichoic acid to the bacterial wall: inhibition by tunicamycin. FEBS Lett. 1976;69(1):75-80. PubMed PMID: 825388.

265.   Soldo B, Lazarevic V, Karamata D. tagO is involved in the synthesis of all anionic cell-wall polymers in Bacillus subtilis 168. Microbiology. 2002;148(Pt 7):2079-87. doi: 10.1099/00221287-148-7-2079. PubMed PMID: 12101296.

266.   Soldo B, Lazarevic V, Pooley HM, Karamata D. Characterization of a Bacillus subtilis thermosensitive teichoic acid-deficient mutant: gene mnaA (yvyH) encodes the UDP-N-acetylglucosamine 2-epimerase. Journal of bacteriology. 2002;184(15):4316-20. PubMed PMID: 12107153; PMCID: PMC135192.

267.   Ashby MN. CaaX converting enzymes. Curr Opin Lipidol. 1998;9(2):99-102. PubMed PMID: 9559265.

268.   Pei J, Grishin NV. Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases. Trends Biochem Sci. 2001;26(5):275-7. PubMed PMID: 11343912.

269.   Neubauer O, Reiffler C, Behrendt L, Eitinger T. Interactions among the A and T Units of an ECF-Type Biotin Transporter Analyzed by Site-Specific Crosslinking. PloS one. 2011;6(12). doi: ARTN e29087

10.1371/journal.pone.0029087. PubMed PMID: WOS:000300674900011.

270.   DeDent A, Bae T, Missiakas DM, Schneewind O. Signal peptides direct surface proteins to two distinct envelope locations of Staphylococcus aureus. Embo Journal. 2008;27(20):2656-68. doi: 10.1038/emboj.2008.185. PubMed PMID: WOS:000260501100003.

271.   Bae T, Schneewind O. The YSIRK-G/S motif of staphylococcal protein A and its role in efficiency of signal peptide processing. Journal of bacteriology. 2003;185(9):2910-9. doi: 10.1228/Jb.185.9.2910-2919.2003. PubMed PMID: WOS:000182459500026.

272.   DeDent AC, McAdow M, Schneewind O. Distribution of protein A on the surface of Staphylococcus aureus. Journal of bacteriology. 2007;189(12):4473-84. doi: 10.1128/JB.00227-07. PubMed PMID: 17416657; PMCID: PMC1913371.

273.   Mazmanian SK, Liu G, Ton-That H, Schneewind O. Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. Science. 1999;285(5428):760-3. PubMed PMID: 10427003.

274.   Perry AM, Ton-That H, Mazmanian SK, Schneewind O. Anchoring of surface proteins to the cell wall of Staphylococcus aureus. III. Lipid II is an in vivo peptidoglycan substrate for sortase-catalyzed surface protein anchoring. The Journal of biological chemistry. 2002;277(18):16241-8. doi: 10.1074/jbc.M109194200. PubMed PMID: 11856734.

275.   Schlag M, Biswas R, Krismer B, Kohler T, Zoll S, Yu W, Schwarz H, Peschel A, Gotz F. Role of staphylococcal wall teichoic acid in targeting the major autolysin Atl. Molecular microbiology. 2010;75(4):864-73. doi: 10.1111/j.1365-2958.2009.07007.x. PubMed PMID: 20105277.

276.   Frankel MB, Hendrickx AP, Missiakas DM, Schneewind O. LytN, a murein hydrolase in the cross-wall compartment of Staphylococcus aureus, is involved in proper bacterial growth and envelope assembly. The Journal of biological chemistry. 2011;286(37):32593-605. doi: 10.1074/jbc.M111.258863. PubMed PMID: 21784864; PMCID: PMC3173183.

277.   Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu LF, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu JH, Shen J. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis. 2016;16(2):161-8. doi: 10.1016/S1473-3099(15)00424-7. PubMed PMID: 26603172.

278.   Friedman L, Alder JD, Silverman JA. Genetic changes that correlate with reduced susceptibility to daptomycin in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2006;50(6):2137-45. doi: 10.1128/AAC.00039-06. PubMed PMID: 16723576; PMCID: PMC1479123.

279.   Yang SJ, Xiong YQ, Dunman PM, Schrenzel J, Francois P, Peschel A, Bayer AS. Regulation of mprF in daptomycin-nonsusceptible Staphylococcus aureus strains. Antimicrobial agents and chemotherapy. 2009;53(6):2636-7. doi: 10.1128/AAC.01415-08. PubMed PMID: 19289517; PMCID: PMC2687189.

280.   Maki H, Yamaguchi T, Murakami K. Cloning and characterization of a gene affecting the methicillin resistance level and the autolysis rate in Staphylococcus aureus. Journal of bacteriology. 1994;176(16):4993-5000. PubMed PMID: 8051012; PMCID: PMC196337.

281.   Torok ME, Chantratita N, Peacock SJ. Bacterial gene loss as a mechanism for gain of antimicrobial resistance. Current opinion in microbiology. 2012;15(5):583-7. doi: 10.1016/j.mib.2012.07.008. PubMed PMID: 23022568; PMCID: PMC3712167.

282.	Gladki A, Kaczanowski S, Szczesny P, Zielenkiewicz P. The evolutionary rate of antibacterial drug targets. BMC Bioinformatics. 2013;14:36. doi: 10.1186/1471-2105-14-36. PubMed PMID: 23374913; PMCID: PMC3598507.

283.	Pulido MR, Garcia-Quintanilla M, Gil-Marques ML, McConnell MJ. Identifying targets for antibiotic development using omics technologies. Drug Discov Today. 2016;21(3):465-72. doi: 10.1016/j.drudis.2015.11.014. PubMed PMID: 26691873.

284.	Nonejuie P, Burkart M, Pogliano K, Pogliano J. Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(40):16169-74. doi: 10.1073/pnas.1311066110. PubMed PMID: 24046367; PMCID: PMC3791758.

285.	Eustice DC, Feldman PA, Slee AM. The mechanism of action of DuP 721, a new antibacterial agent: effects on macromolecular synthesis. Biochem Biophys Res Commun. 1988;150(3):965-71. PubMed PMID: 2449210.

286.	Hutter B, Schaab C, Albrecht S, Borgmann M, Brunner NA, Freiberg C, Ziegelbauer K, Rock CO, Ivanov I, Loferer H. Prediction of mechanisms of action of antibacterial compounds by gene expression profiling. Antimicrobial agents and chemotherapy. 2004;48(8):2838-44. doi: 10.1128/AAC.48.8.2838-2844.2004. PubMed PMID: 15273089; PMCID: PMC478524.

287.	Bandow JE, Brotz H, Leichert LI, Labischinski H, Hecker M. Proteomic approach to understanding antibiotic action. Antimicrobial agents and chemotherapy. 2003;47(3):948-55. PubMed PMID: 12604526; PMCID: PMC149304.

288.	Torrent M, Andreu D, Nogues VM, Boix E. Connecting peptide physicochemical and antimicrobial properties by a rational prediction model. PloS one. 2011;6(2):e16968. doi: 10.1371/journal.pone.0016968. PubMed PMID: 21347392; PMCID: PMC3036733.

289.	Johnston CW, Skinnider MA, Dejong CA, Rees PN, Chen GM, Walker CG, French S, Brown ED, Berdy J, Liu DY, Magarvey NA. Assembly and clustering of natural antibiotics guides target identification. Nature chemical biology. 2016;12(4):233-9. doi: 10.1038/nchembio.2018. PubMed PMID: 26829473.

290.	Wildenhain J, Spitzer M, Dolma S, Jarvik N, White R, Roy M, Griffiths E, Bellows DS, Wright GD, Tyers M. Prediction of synergism from chemical-genetic Interactions by machine learning. Cell Systems. 2015;1(6):383-95. doi: 10.1016/j.cels.2015.12.003.

291.	Klein BA, Tenorio EL, Lazinski DW, Camilli A, Duncan MJ, Hu LT. Identification of essential genes of the periodontal pathogen Porphyromonas gingivalis. BMC genomics. 2012;13:578. doi: 10.1186/1471-2164-13-578. PubMed PMID: 23114059; PMCID: PMC3547785.

292.	Drlica K. Mechanism of fluoroquinolone action. Current opinion in microbiology. 1999;2(5):504-8. PubMed PMID: 10508721.

293.	Kotra LP, Haddad J, Mobashery S. Aminoglycosides: perspectives on mechanisms of action and resistance and strategies to counter resistance. Antimicrobial agents and chemotherapy. 2000;44(12):3249-56. PubMed PMID: 11083623; PMCID: PMC90188.

294.   Mates SM, Eisenberg ES, Mandel LJ, Patel L, Kaback HR, Miller MH. Membrane potential and gentamicin uptake in Staphylococcus aureus. Proceedings of the National Academy of Sciences of the United States of America. 1982;79(21):6693-7. PubMed PMID: 6959147; PMCID: PMC347195.

295.   Bryan LE, Van Den Elzen HM. Effects of membrane-energy mutations and cations on streptomycin and gentamicin accumulation by bacteria: a model for entry of streptomycin and gentamicin in susceptible and resistant bacteria. Antimicrobial agents and chemotherapy. 1977;12(2):163-77. PubMed PMID: 143238; PMCID: PMC429880.

296.   Morikawa K, Maruyama A, Inose Y, Higashide M, Hayashi H, Ohta T. Overexpression of sigma factor, sigma(B), urges Staphylococcus aureus to thicken the cell wall and to resist beta-lactams. Biochem Biophys Res Commun. 2001;288(2):385-9. doi: 10.1006/bbrc.2001.5774. PubMed PMID: 11606054.

297.   Kuroda M, Kuroda H, Oshima T, Takeuchi F, Mori H, Hiramatsu K. Two-component system VraSR positively modulates the regulation of cell-wall biosynthesis pathway in Staphylococcus aureus. Molecular microbiology. 2003;49(3):807-21. Epub 2003/07/17. PubMed PMID: 12864861.

298.   Neyfakh AA, Borsch CM, Kaatz GW. Fluoroquinolone resistance protein NorA of Staphylococcus aureus is a multidrug efflux transporter. Antimicrobial agents and chemotherapy. 1993;37(1):128-9. PubMed PMID: 8431010; PMCID: PMC187619.

299.   Young IG, Rogers BL, Campbell HD, Jaworowski A, Shaw DC. Nucleotide sequence coding for the respiratory NADH dehydrogenase of Escherichia coli. UUG initiation codon. Eur J Biochem. 1981;116(1):165-70. PubMed PMID: 6265208.

300.   Komatsuzawa H, Sugai M, Ohta K, Fujiwara T, Nakashima S, Suzuki J, Lee CY, Suginaka H. Cloning and characterization of the fmt gene which affects the methicillin resistance level and autolysis in the presence of triton X-100 in methicillin-resistant Staphylococcus aureus. Antimicrobial agents and chemotherapy. 1997;41(11):2355-61. Epub 1997/11/26. PubMed PMID: 9371333; PMCID: 164128.

301.   Qamar A, Golemi-Kotra D. Dual roles of FmtA in Staphylococcus aureus cell wall biosynthesis and autolysis. Antimicrobial agents and chemotherapy. 2012;56(7):3797-805. doi: 10.1128/AAC.00187-12. PubMed PMID: 22564846; PMCID: PMC3393393.

302.   Rahman MM, Hunter HN, Prova S, Verma V, Qamar A, Golemi-Kotra D. The Staphylococcus aureus Methicillin Resistance Factor FmtA Is a d-Amino Esterase That Acts on Teichoic Acids. mBio. 2015;7(1). doi: 10.1128/mBio.02070-15. PubMed PMID: 26861022; PMCID: PMC4752606.

303.   Falord M, Karimova G, Hiron A, Msadek T. GraXSR proteins interact with the VraFG ABC transporter to form a five-component system required for cationic antimicrobial peptide sensing and resistance in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2012;56(2):1047-58. Epub 2011/11/30. doi: 10.1128/AAC.05054-11. PubMed PMID: 22123691; PMCID: 3264281.

304.   Falord M, Mader U, Hiron A, Debarbouille M, Msadek T. Investigation of the Staphylococcus aureus GraSR regulon reveals novel links to virulence, stress response and cell wall signal transduction pathways. PloS one. 2011;6(7):e21323. doi: 10.1371/journal.pone.0021323. PubMed PMID: 21765893; PMCID: PMC3128592.

305.   Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. Annu Rev Biochem. 2000;69:183-215. doi: 10.1146/annurev.biochem.69.1.183. PubMed PMID: 10966457.

306.   Arvidson S, Tegmark K. Regulation of virulence determinants in Staphylococcus aureus. International journal of medical microbiology : IJMM. 2001;291(2):159-70. doi: 10.1078/1438-4221-00112. PubMed PMID: 11437338.

307.   George EA, Muir TW. Molecular mechanisms of agr quorum sensing in virulent staphylococci. Chembiochem : a European journal of chemical biology. 2007;8(8):847-55. doi: 10.1002/cbic.200700023. PubMed PMID: 17457814.

308.   Singh VK, Carlos MR, Singh K. Physiological significance of the peptidoglycan hydrolase, LytM, in Staphylococcus aureus. FEMS Microbiol Lett. 2010;311(2):167-75. doi: 10.1111/j.1574-6968.2010.02087.x. PubMed PMID: 20738399; PMCID: PMC2944916.

309.   Lioliou E, Fechter P, Caldelari I, Jester BC, Dubrac S, Helfer AC, Boisset S, Vandenesch F, Romby P, Geissmann T. Various checkpoints prevent the synthesis of Staphylococcus aureus peptidoglycan hydrolase LytM in the stationary growth phase. RNA Biol. 2016:1-14. doi: 10.1080/15476286.2016.1153209. PubMed PMID: 26901414.

310.   Piriz Duran S, Kayser FH, Berger-Bachi B. Impact of sar and agr on methicillin resistance in Staphylococcus aureus. FEMS Microbiol Lett. 1996;141(2-3):255-60. PubMed PMID: 8768531.

311.   Ernst CM, Staubitz P, Mishra NN, Yang SJ, Hornig G, Kalbacher H, Bayer AS, Kraus D, Peschel A. The bacterial defensin resistance protein MprF consists of separable domains for lipid lysinylation and antimicrobial peptide repulsion. PLoS Pathog. 2009;5(11):e1000660. doi: 10.1371/journal.ppat.1000660. PubMed PMID: 19915718; PMCID: PMC2774229.

312.   Neuhaus FC, Baddiley J. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. Microbiol Mol Biol Rev. 2003;67(4):686-723. PubMed PMID: 14665680; PMCID: PMC309049.

313.   Peschel A, Otto M, Jack RW, Kalbacher H, Jung G, Gotz F. Inactivation of the dlt operon in Staphylococcus aureus confers sensitivity to defensins, protegrins, and other antimicrobial peptides. The Journal of biological chemistry. 1999;274(13):8405-10. PubMed PMID: 10085071.

314.   Biswas R, Martinez RE, Gohring N, Schlag M, Josten M, Xia G, Hegler F, Gekeler C, Gleske AK, Gotz F, Sahl HG, Kappler A, Peschel A. Proton-binding capacity of Staphylococcus aureus wall teichoic acid and its role in controlling autolysin activity. PloS one. 2012;7(7):e41415. doi: 10.1371/journal.pone.0041415. PubMed PMID: 22911791; PMCID: PMC3402425.

315.   Gross M, Cramton SE, Gotz F, Peschel A. Key role of teichoic acid net charge in Staphylococcus aureus colonization of artificial surfaces. Infect Immun. 2001;69(5):3423-6. doi: 10.1128/IAI.69.5.3423-3426.2001. PubMed PMID: 11292767; PMCID: PMC98303.

316.   Lunsford RD, Roble AG. comYA, a gene similar to comGA of Bacillus subtilis, is essential for competence-factor-dependent DNA transformation in Streptococcus gordonii. Journal of bacteriology. 1997;179(10):3122-6. PubMed PMID: 9150204; PMCID: PMC179087.

317.   Wadstrom T. Chitinase activity and substrate specificity of endo- -N-acetyl-glucosaminidase of Staphylococcus aureus, strain M18. Acta Chem Scand. 1971;25(5):1807-12. PubMed PMID: 5099106.

318.   Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825-30. PubMed PMID: WOS:000298103200003.

319.   Kaito C, Kurokawa K, Matsumoto Y, Terao Y, Kawabata S, Hamada S, Sekimizu K. Silkworm pathogenic bacteria infection model for identification of novel virulence genes. Molecular microbiology. 2005;56(4):934-44. doi: 10.1111/j.1365-2958.2005.04596.x. PubMed PMID: WOS:000229071100007.

320.   Ikuo M, Kaito C, Sekimizu K. The cvfC operon of Staphylococcus aureus contributes to virulence via expression of the thyA gene. Microb Pathog. 2010;49(1-2):1-7. doi: 10.1016/j.micpath.2010.03.010. PubMed PMID: 20347953.

321.   Stone KJ, Stromiger JL. Mechanism of action of bacitracin: complexation with metal ion and C 55 -isoprenyl pyrophosphate. Proceedings of the National Academy of Sciences of the United States of America. 1971;68(12):3223-7. Epub 1971/12/01. PubMed PMID: 4332017; PMCID: 389626.

322.   Wallhausser KH, Nesemann G, Prave P, Steigler A. Moenomycin, a new antibiotic. I. Fermentation and isolation. Antimicrob Agents Chemother (Bethesda). 1965;5:734-6. PubMed PMID: 5883490.

323.   Kahan FM, Kahan JS, Cassidy PJ, Kropp H. The mechanism of action of fosfomycin (phosphonomycin). Annals of the New York Academy of Sciences. 1974;235(0):364-86. Epub 1974/05/10. PubMed PMID: 4605290.

324.   Suzuki T, Swoboda JG, Campbell J, Walker S, Gilmore MS. In vitro antimicrobial activity of wall teichoic acid biosynthesis inhibitors against Staphylococcus aureus isolates. Antimicrobial agents and chemotherapy. 2011;55(2):767-74. Epub 2010/11/26. doi: 10.1128/AAC.00879-10. PubMed PMID: 21098254; PMCID: 3028815.

325.   Hughes J, Mellows G. Inhibition of isoleucyl-transfer ribonucleic acid synthetase in Escherichia coli by pseudomonic acid. Biochem J. 1978;176(1):305-18. PubMed PMID: 365175; PMCID: PMC1186229.

326.   Wehrli W, Nuesch J, Knusel F, Staehelin M. Action of rifamycins on RNA polymerase. Biochimica et biophysica acta. 1968;157(1):215-7. PubMed PMID: 4171079.

327.   Schmitt J, Joost I, Skaar EP, Herrmann M, Bischoff M. Haemin represses the haemolytic activity of Staphylococcus aureus in an Sae-dependent manner. Microbiology. 2012;158(Pt 10):2619-31. doi: 10.1099/mic.0.060129-0. PubMed PMID: 22859613; PMCID: PMC4083625.

328.   Tegmark K, Morfeldt E, Arvidson S. Regulation of agr-dependent virulence genes in Staphylococcus aureus by RNAIII from coagulase-negative staphylococci. Journal of bacteriology. 1998;180(12):3181-6. PubMed PMID: 9620969; PMCID: PMC107820.

329.   Giraudo AT, Cheung AL, Nagel R. The sae locus of Staphylococcus aureus controls exoprotein synthesis at the transcriptional level. Arch Microbiol. 1997;168(1):53-8. PubMed PMID: 9211714.

330.   Novick RP, Jiang D. The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. Microbiology. 2003;149(Pt 10):2709-17. doi: 10.1099/mic.0.26575-0. PubMed PMID: 14523104.

331.   Rogasch K, Ruhmling V, Pane-Farre J, Hoper D, Weinberg C, Fuchs S, Schmudde M, Broker BM, Wolz C, Hecker M, Engelmann S. Influence of the two-component system SaeRS on global gene expression in two different Staphylococcus aureus strains. Journal of bacteriology. 2006;188(22):7742-58. doi: 10.1128/JB.00555-06. PubMed PMID: 17079681; PMCID: PMC1636327.

332.   Malachowa N, Whitney AR, Kobayashi SD, Sturdevant DE, Kennedy AD, Braughton KR, Shabb DW, Diep BA, Chambers HF, Otto M, DeLeo FR. Global changes in Staphylococcus aureus gene expression in human blood. PloS one. 2011;6(4):e18617. doi: 10.1371/journal.pone.0018617. PubMed PMID: 21525981; PMCID: PMC3078114.

333.   Anzaldi LL, Skaar EP. Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. Infect Immun. 2010;78(12):4977-89. doi: 10.1128/IAI.00613-10. PubMed PMID: 20679437; PMCID: PMC2981329.

334.   Reniere ML, Torres VJ, Skaar EP. Intracellular metalloporphyrin metabolism in Staphylococcus aureus. Biometals. 2007;20(3-4):333-45. doi: 10.1007/s10534-006-9032-0. PubMed PMID: 17387580.

335.   El Ghachi M, Bouhss A, Blanot D, Mengin-Lecreulx D. The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. The Journal of biological chemistry. 2004;279(29):30106-13. Epub 2004/05/13. doi: 10.1074/jbc.M401701200. PubMed PMID: 15138271.

336.   Bayer AS, Mishra NN, Sakoulas G, Nonejuie P, Nast CC, Pogliano J, Chen KT, Ellison SN, Yeaman MR, Yang SJ. Heterogeneity of mprF Sequences in Methicillin-Resistant Staphylococcus aureus Clinical Isolates: Role in Cross-Resistance between Daptomycin and Host Defense Antimicrobial Peptides. Antimicrobial agents and chemotherapy. 2014. Epub 2014/10/08. doi: 10.1128/AAC.03422-14. PubMed PMID: 25288091.

337.   Halouska S, Fenton RJ, Zinniel DK, Marshall DD, Barletta RG, Powers R. Metabolomics analysis identifies d-Alanine-d-Alanine ligase as the primary lethal target of d-Cycloserine in mycobacteria. J Proteome Res. 2014;13(2):1065-76. doi: 10.1021/pr4010579. PubMed PMID: 24303782; PMCID: PMC3975674.

338.   Sham LT, Butler EK, Lebar MD, Kahne D, Bernhardt TG, Ruiz N. Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. Science. 2014;345(6193):220-2. doi: 10.1126/science.1254522. PubMed PMID: 25013077; PMCID: PMC4163187.

339.   Fang X, Tiyanont K, Zhang Y, Wanner J, Boger D, Walker S. The mechanism of action of ramoplanin and enduracidin. Molecular bioSystems. 2006;2(1):69-76. Epub 2006/08/02. doi: 10.1039/b515328j. PubMed PMID: 16880924.

340.   Siewert G, Strominger JL. Bacitracin: an inhibitor of the dephosphorylation of lipid pyrophosphate, an intermediate in the biosynthesis of the peptidoglycan of bacterial cell walls. Proc Natl Acad Sci U S A. 1967;57(3):767-73. PubMed PMID: 16591529; PMCID: PMC335574.

341.   Teuber M, Bader J. Action of polymyxin B on bacterial membranes. Binding capacities for polymyxin B of inner and outer membranes isolated from Salmonella typhimurium G30. Arch Microbiol. 1976;109(1-2):51-8. PubMed PMID: 183617.

342.   Kanamori M, Kamata H, Yagisawa H, Hirata H. Overexpression of the alanine carrier protein gene from thermophilic bacterium PS3 in Escherichia coli. J Biochem. 1999;125(3):454-9. PubMed PMID: 10050032.

343.   Kohanski MA, Dwyer DJ, Wierzbowski J, Cottarel G, Collins JJ. Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. Cell. 2008;135(4):679-90. doi: 10.1016/j.cell.2008.09.038. PubMed PMID: 19013277; PMCID: PMC2684502.

344.   Campbell J, Singh AK, Swoboda JG, Gilmore MS, Wilkinson BJ, Walker S. An antibiotic that inhibits a late step in wall teichoic acid biosynthesis induces the cell wall stress stimulon in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2012;56(4):1810-20. doi: 10.1128/AAC.05938-11. PubMed PMID: 22290958; PMCID: PMC3318382.

345.   Huber J, Donald RG, Lee SH, Jarantow LW, Salvatore MJ, Meng X, Painter R, Onishi RH, Occi J, Dorso K, Young K, Park YW, Skwish S, Szymonifka MJ, Waddell TS, Miesel L, Phillips JW, Roemer T. Chemical genetic identification of peptidoglycan inhibitors potentiating carbapenem activity against methicillin-resistant Staphylococcus aureus. Chem Biol. 2009;16(8):837-48. doi: 10.1016/j.chembiol.2009.05.012. PubMed PMID: 19716474.

346.   Lee K, Campbell J, Swoboda JG, Cuny GD, Walker S. Development of improved inhibitors of wall teichoic acid biosynthesis with potent activity against Staphylococcus aureus. Bioorg Med Chem Lett. 2010;20(5):1767-70. doi: 10.1016/j.bmcl.2010.01.036. PubMed PMID: 20138521; PMCID: PMC2844852.

347.   Champness JN, Stammers DK, Beddell CR. Crystallographic investigation of the cooperative interaction between trimethoprim, reduced cofactor and dihydrofolate reductase. FEBS letters. 1986;199(1):61-7. PubMed PMID: 3082676.

348.   Zhang W, Li Y, Qian G, Wang Y, Chen H, Li YZ, Liu F, Shen Y, Du L. Identification and characterization of the anti-methicillin-resistant Staphylococcus aureus WAP-8294A2 biosynthetic gene cluster from Lysobacter enzymogenes OH11. Antimicrobial agents and chemotherapy. 2011;55(12):5581-9. doi: 10.1128/AAC.05370-11. PubMed PMID: 21930890; PMCID: PMC3232812.

349.   Kato A, Hirata H, Ohashi Y, Fujii K, Mori K, Harada K. A new anti-MRSA antibiotic complex, WAP-8294A II. Structure characterization of minor components by ESI LCMS and MS/MS. The Journal of antibiotics. 2011;64(5):373-9. doi: 10.1038/ja.2011.9. PubMed PMID: 21326252.

350.   Kato A, Nakaya S, Kokubo N, Aiba Y, Ohashi Y, Hirata H, Fujii K, Harada K. A new anti-MRSA antibiotic complex, WAP-8294A. I. Taxonomy, isolation and biological activities. The Journal of antibiotics. 1998;51(10):929-35. PubMed PMID: 9917006.

351.   Miyazaki E, Chen JM, Ko C, Bishai WR. The Staphylococcus aureus rsbW (orf159) gene encodes an anti-sigma factor of SigB. Journal of bacteriology. 1999;181(9):2846-51. PubMed PMID: 10217777; PMCID: PMC93728.

352.   Palma M, Cheung AL. sigma(B) activity in Staphylococcus aureus is controlled by RsbU and an additional factor(s) during bacterial growth. Infect Immun. 2001;69(12):7858-65. doi: 10.1128/IAI.69.12.7858-7865.2001. PubMed PMID: 11705968; PMCID: PMC98882.

353.   Kullik II, Giachino P. The alternative sigma factor sigmaB in Staphylococcus aureus: regulation of the sigB operon in response to growth phase and heat shock. Arch Microbiol. 1997;167(2/3):151-9. PubMed PMID: 9042755.

354.   Tuchscherr L, Bischoff M, Lattar SM, Noto Llana M, Pfortner H, Niemann S, Geraci J, Van de Vyver H, Fraunholz MJ, Cheung AL, Herrmann M, Volker U, Sordelli DO, Peters G, Loffler B. Sigma Factor SigB Is Crucial to Mediate Staphylococcus aureus Adaptation during Chronic Infections. PLoS Pathog. 2015;11(4):e1004870. doi: 10.1371/journal.ppat.1004870. PubMed PMID: 25923704; PMCID: PMC4414502.

355.   Bischoff M, Berger-Bachi B. Teicoplanin stress-selected mutations increasing sigma(B) activity in Staphylococcus aureus. Antimicrobial agents and chemotherapy. 2001;45(6):1714-20. doi: 10.1128/AAC.45.6.1714-1720.2001. PubMed PMID: 11353616; PMCID: PMC90536.

356.   Pelletier LL, Jr., Richardson M, Feist M. Virulent gentamicin-induced small colony variants of Staphylococcus aureus. J Lab Clin Med. 1979;94(2):324-34. PubMed PMID: 458250.

357.   Xu Q, Bateman A, Finn RD, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Carlton D, Chen C, Chiu HJ, Chiu M, Clayton T, Das D, Deller MC, Duan L, Ellrott K, Ernst D, Farr CL, Feuerhelm J, Grant JC, Grzechnik A, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Krishna SS, Kumar A, Marciano D, McMullan D, Miller MD, Morse AT, Nigoghossian E, Nopakun A, Okach L, Puckett C, Reyes R, Rife CL, Sefcovic N, Tien HJ, Trame CB, van den Bedem H, Weekes D, Wooten T, Hodgson KO, Wooley J, Elsliger MA, Deacon AM, Godzik A, Lesley SA, Wilson IA. Bacterial pleckstrin homology domains: a prokaryotic origin for the PH domain. J Mol Biol. 2010;396(1):31-46. doi: 10.1016/j.jmb.2009.11.006. PubMed PMID: 19913036; PMCID: PMC2817789.

358.   Yu JW, Mendrola JM, Audhya A, Singh S, Keleti D, DeWald DB, Murray D, Emr SD, Lemmon MA. Genome-wide analysis of membrane targeting by S. cerevisiae pleckstrin homology domains. Mol Cell. 2004;13(5):677-88. PubMed PMID: 15023338.

359.   Wang DS, Shaw G. The association of the C-terminal region of beta I sigma II spectrin to brain membranes is mediated by a PH domain, does not require membrane proteins, and coincides with a inositol-1,4,5 triphosphate binding site. Biochem Biophys Res Commun. 1995;217(2):608-15. PubMed PMID: 7503742.

360.   Kuhn S, Slavetinsky CJ, Peschel A. Synthesis and function of phospholipids in Staphylococcus aureus. International journal of medical microbiology : IJMM. 2015;305(2):196-202. doi: 10.1016/j.ijmm.2014.12.016. PubMed PMID: 25595024.

361.   Hachmann AB, Sevim E, Gaballa A, Popham DL, Antelmann H, Helmann JD. Reduction in membrane phosphatidylglycerol content leads to daptomycin resistance in Bacillus subtilis. Antimicrobial agents and chemotherapy. 2011;55(9):4326-37. doi: 10.1128/AAC.01819-10. PubMed PMID: 21709092; PMCID: PMC3165287.

Appendices

A. Methods

Strains and plasmids (Chapter 2-5)

Table of strains and plasmids used can be found in Appendix B. Original strains and plasmids for the phage-based transposition system were obtained from Merck Research Laboratories. Plasmids were first introduced into the restriction negative S. aureus strain RN4220 and then moved into restriction competent HG003 strains by electroporation. Gene deletion strains were made using the E. coli-S. aureus shuttle vector pKFC. Stellar competent cells, pUC19, and the In-Fusion cloning kit were purchased from Clontech. Cultures of S. aureus were routinely grown in tryptic soy broth (TSB) at the indicated temperature, with erythromycin (5 or 10 µg/mL), chloramphenicol (10 µg/mL), or kanamycin/neomycin (25 µg/mL each) when required. Phage lysates were prepared in TSB top agar and titered as has been previously described. Overexpression strains were made by cloning the gene of interest into the pLOW plasmid. Mutants from the Nebraska transposon library were validated using PCR of the gene of interest and were transduced into the parent strain before experimentation.

Creation of the transposon library (Chapter 2 and 3)

To create the transposon library, phage lysates were made using the six donor strains harboring plasmids encoding transposon cassettes (pTM239-pTM244), where one in three progeny phage carries the transposon plasmid. Recipient strains (TM231 and the negative control TM232) were grown to late exponential phase, and resuspended in supplemented glucose minimal media (SGMM) (10 mM glucose, 2 mM MgCl2, 3.5 mM CaCl2, 0.1% CAS amino acids, 0.5% NaCl, 10 mM MES, pH 6.8). For each transposon construct, phage lysate was mixed with 25 ml recipient strain in SGMM (multiplicity of infection = 5). Solutions were

incubated at room temperature overnight. The next day, cells were pelleted and washed three times with an equal volume of TSB before being allowed to recover with shaking for two hours at 30°C. Approximately 108-109 cfu were spread per 150 × 15 mm petri dish containing TSB agar with 5 µg/ml erythromycin to select for transposon insertion mutants. The number of spontaneously ermR colonies was less than 1% in the control strain TM232 that does not express functional transposase. Plates were incubated at 30°C for two days, before being scraped (1-2 million total colonies) and pooled into a single 100 mL suspension. Cells were pelleted and washed three times with an equal volume of TSB, before being resuspended in TSB-glycerol (12.5% v/v), aliquoted, and stored at 80°C. Additional details for the preparation and sequencing of the transposon library can be found in Appendix C.

Assessment of next-generation sequencing library composition (Chapter 2)

A linearized pUC19 vector was made by inverse PCR using primers (Tm179-Tm180) containing 5'-overhangs homologous to the termini of the P5 and P7 Illumina sequences. The 2.6 kb PCR product was agarose gel purified, ligated to an aliquot of the NGS insert library using the Clontech In-Fusion system as directed by the manufacturer, transformed into chemically competent E. coli cells, and selected on LB carbenicillin (100 µg/mL) plates. Plasmids were isolated from 2 mL cultures of separate colonies and the DNA inserts sequenced. At least 10 colonies were sequenced to confirm insert sequence diversity and lack of plasmid-transposon junction DNA.

Next-generation sequencing Galaxy analysis (Chapter 2-5)

The DNA concentration was determined using the Quant-IT™ PicoGreen kit from Invitrogen, and the sample was diluted to 10 nM in Buffer EB. Six samples with different

indexing barcodes were multiplexed together, and samples were sequenced on a Hi-Seq2000 or Hi-Seq2500 for 100 cycles with 40% ΦX174 spiked in to the sequencing reaction.

Sequencing data were uploaded to the Tufts Galaxy Service hosted by TUCF Genomics (http://tucf-genomics.tufts.edu/). The first four base pairs of the 8 bp indexing barcode were deleted since these data were occasionally of low quality. Multiplexed sample data was separated by the remaining 4 bp of the indexing barcode, and then again by the 3 bp transposon cassette barcode. Read ends were trimmed to leave the 16 bp of genomic sequence immediately flanking the TA dinucleotide insertion site and filtered to keep high-quality reads (>90% of bases with quality score >20). Remaining reads were mapped to the S. aureus NCTC8325 genome using Bowtie. Hopcounts (the number of reads that map to a given site in the genome) were determined and these files were downloaded for subsequent statistical analysis. For details on custom analyses, see Appendix H

Growth curves (Chapter 2 - 4)

Growth curves were performed by diluting overnight cultures to an OD600 of 0.01 in 25 mL Tryptic Soy Broth. Strains were grown with shaking in a 43°C or 30°C water bath. The OD of the cultures was measured every 10 (43°C) or every 20 (30°C) minutes in an Ultrospec 10 cell density meter until stationary phase.

96-well growth curves (Chapter 3 and 5)

Overnight cultures, were diluted to an OD_{600} = 0.001. 150µl of diluted culture was mixed with 1.5µl of antibiotic at 100X final concentration and mixed in a well of the 96 well plate. Growth curves were performed in 96 well plates at 37°C with shaking, measuring the OD_{600} every 15 or 20min.. A Molecular devices plate reader was used for these experiments.

Treatment of MRSA transposon libraries with β-lactams and sequencing (Chapter 3)

Transposon libraries were treated in 50mL cultures of TSB at 37°C. We inoculated with 1,000,000 cfus of the transposon library and added the appropriate antibiotic. Then cultures were grown, shaking, until the cultures grew to OD between 1 and 2. Cells were spun down, and the DNA was isolated and prepared for Tn-Seq as described in Appendix C.

Library 1 antibiotic treatment (Chapter 4)

Library 1 was constructed by transformation of a temperature sensitive plasmid. A 100 µl aliquot of this initial S. aureus HG003 transposon library freezer stock, containing 10^8cfu, was used to inoculate 100 ml of MH cation adjusted broth and incubated for 15 h at 37°C with shaking at 200 rpm. A 10 µl aliquot (10^6cfu) of this input culture was then inoculated into a final volume of 200 µl in a 96-well plate broth microdilution format and incubated at 37°C for 8 h [approximately 5.5 generations (5 x 10^7cfu/200 µl)]. The 1/2x, 1/4x and 1/8x MIC wells for the library pool were determined based on the MIC of a small mutant pool (consisting of ten innocuous transposon mutants). This small pool was used to determine MICs in order to compensate for potential resistant mutants in the library pool. The chosen wells (1/2x, 1/4x and 1/8x) were then subcultured (3 x 10^5cfu) into a second iteration of serial dilutions of antibiotics as described above and incubated for 15 h at 37°C [approximately 9 generations (2 x 10^8cfu/200 µl)]. The 1/2x, 1/4x and 1/8x MIC wells were determined based on the small pool and these wells were transferred to 10 ml of BHI broth, incubated for 4 h at 37°C with shaking at 200 rpm. Biological replicates were conducted for each growth condition. Genomic DNA was harvested using DNeasy Blood and Tissue kit (Qiagen, Valencia, CA) following the manufacturer's instructions. Library1 was prepared using the shearing method. Illumina sequencing was completed by Harvard Biopolymers Facility or Tufts Genomic DNA sequencing core facility.

Library 2 antibiotic treatment (Chapter 4 and 5)

Library 2 was constructed using phage-based transposition and six different transposon constructs as previously described. TSB supplemented with 25mg/L Ca2+ and 12.5mg/L Mg2+ was used for all antibiotics except for oxacillin, which was tested using cation-adjusted Mueller-Hinton Broth. For all antibiotics, an untreated control was prepared in the same media as was used for the tested antibiotic. A stock of the complete library was thawed and diluted to 0.2 OD600 and grown to an OD 600 of ~0.4 to allow for a minimal stimulation of growth prior to treatment. The culture was then diluted to 4 x 10^5cfu/mL and added to 1mL of media with 2x the desired concentration of the antibiotic, to give a final starting inoculum of 2x10^5cfu/mL in 2mL culture volumes. Samples were grown at 37°C and harvested when they reached stationary phase. The concentrations of antibiotic tested were 2x, 1x, 0.5x and 0.25 x the MIC of the antibiotic. Samples were prepared for NGS as described in Appendix C.

Plate dilution spotting (Chapter 3)

Agar plates were prepared with TSB and the desired concentration of the compound/antibiotic of interest. Overnight cultures of mutants as well as WT strains were diluted 1:100 in fresh TSB and grown to an OD of 0.5. They were then diluted serially by 10-fold, spotted on to agar plate and incubated at 37°C overnight. Photos of plates were taken 18-36 hours after plating.

Plate dilution spotting (Chapter 4)

Identified hits *ndh, fmtA, SAOUHSC_01025* and *SAOUHSC_01050* were validated using transposon mutants from the Nebraska library in background USA300_FPR3757.Deletion mutants in *dltA,graR,* and *mprF* were tested in MSSA strain Newman. Agar plates were prepared with TSB supplemented with Ca2+ (25mg/L) and Mg2+ (12.5mg/L) and a sub-MIC

concentration of the 6 antibiotics. Overnight cultures of mutants were diluted 1:100 in fresh TSB and grown to an OD of 1. They were then diluted serially by 10fold, spotted on to agar plate and incubated at 37°C overnight. The concentration of the antibiotic at which WT was severely inhibited showing growth in only the highest 1 or 2 dilutions on agar plates under these conditions was determined. This was considered to be the MIC under this condition. The spot dilution assays were then set up using three different antibiotic concentrations. The concentration closest to the MIC at which WT was at most 3 logs more depleted than non-antibiotic control was used to calculate fitness. This concentration was used so that reduced fitness of any mutants could be observed. The exception to this was when the MIC concentration for gentamicin was used to evaluate the resistance of inactivation mutants in the oxidative phosphorylation pathway. Controls plates with no antibiotic were set up for all strains assayed and under these conditions, mutants and WT showed equal levels of growth (not shown). Fitness was assessed by determining the highest dilution for which growth was observed for a mutant and the WT strain. The highest dilution showing full growth for the mutant was then divided by the highest dilution showing full growth for the WT to calculate its fitness compared to WT. These were plotted on a log scale. Those spots that showed hazy growth indicative of cell lysis, those that showed mixed populations of colonies of different sizes suggesting the possibility of suppressors and reduced fitness relative to spots with homogenous colonies, and those that had fewer than 10 individual colonies were not regarded as full growth.

Co-immunoprecipitation experiments using LyrA (Chapter 3)

Pull downs were performed by expressing LyrA linked to c-myc in a *ΔlyrA.* Cells were grown and then DSP crosslinking reagent was added. Cells were disrupted and membrane proteins solubilized. Then, the solubilized membrane proteins were run over a column with the tag receptor. Tagged LyrA as well as anything it is interacting with binds to the column. All other

proteins are washed away, and then LyrA and anything it is interaction with is eluted. These proteins are boiled, run on SDS-PAGE gels. Bands are cut out and digested with Trypsin. The identity of these peptides is determined using LC-MS.

Transmission electron microscopy (Chapter 3)

Overnight cultures of the strains of interest were diluted to an OD=0.05 in 2mL TSB and grown to early exponential phase (OD~0.3). Cells were spun down 8000g 5min and resuspended in 100µl water and 100µl glutaraldehyde-formaledyle in sodium cacodylate buffer pH7.4 fixative provided by the Harvard Electron Microscopy core facility who also embedded the samples in resins. Samples were viewed using the JEOL 1200EX – 80kV electron microscope.

Transmission electron microscopy with gold nanoparticles(Chapter 3)

Overnight cultures of cells were dilution to an OD=0.5 in 4mL of TSB, and grown to early exponential phase (OD~0.3). Cells were aliquoted into 2 x 2mL and spun down 8000g 5min and washed twice with 1mL PBS pH 7.4. Cells were resuspended in 1mL PBS + 0.2mg/ml trypsin and incubated at 37C for 1 hour to remove proteins attached to the cell surface. Cells were spun down again and washed twice with 1mL PBS pH 7.4. One set of tubes was kept aside as a control, while the other was resuspended in 2mL TSB + 1mM PMSF (a protease inhibitor), transferred to a 15mL culture tube, and grown at 37°C, shaking for 10min to allow surface protein synthesis and export again. After this, the cells are spun down and washed twice with PBS pH 7.4 again. Then, they were resuspended in 1mL PBS + 1:2000 goat anti-protein A and incubated at 37°C for 20 minutes. After another cycle of spinning down and washing with PBS pH 7.4, cells were resuspend in 50µl + 1:20 rabbit anti-goat antibody conjugated to 10nm gold nanobeads, incubated at room temperature for 20 minutes, and spun down. After washing twice with PBS pH 7.4, they were resuspended in 50µl PBS. 50µl of the

same fixative described above was added, the Harvard Electron Microscopy Core embedded the samples in resins, and the samples were viewed using the JEOL 1200EX – 80kV electron microscope. Antibodies were purchased from Abcam, product numbers ab181627 (Anti-protein A) and ab27245 (Rabbit anti-goat Gold).

Identification of antibiotic resistance factors (Chapter 4)

We identified datasets from Library 2 where reads mapped to 25-40%of the TA sites hit in the untreated control (with the exception of vancomycin treatment which hit 67% of the TA sites hit in the untreated control). These were processed for further analysis. This percent decrease was chosen such that we could identify genes with an increase and decrease in number of reads mapping to them. Library 2 contains transposon constructs with outward-facing promoters that can upregulate proximal genes in addition to the traditional construct which can only insert into and inactivate genes. For these experiments, we only considered data from the inactivation constructs. Data from biological replicates was combined, and before comparing the number of reads/gene using the Mann-Whitney U test, the experimental data was normalized to the control using simulation-based resampling. Then, data for each antibiotic treatment from each of the Library 1 experiments was combined with the data from the Library 2 experiments using the geometric mean of the ratios and Fisher's method for combining corrected p-values. Top hits were identified by first filtering for genes with a p-value less than 0.05, then by increasing the ratio cut-off by integers until less than or equal to 20 genes were left. See Appendix H for scripts.

Triton X-100 induced lysis assay (Chapter 4)

Overnight cultures of WT and mutants were diluted 1:100 in fresh TSB and grown to an OD of 0.6. The cultures were harvested, washed with cold, sterile water and re-suspended in

0.05M Tris-HCl, pH 7.2, with or without 0.05% Triton X-100. Samples were incubated at 37°C

and OD600 was measured every 20min for 3 hours. OD was normalized to the initial

measurement for each sample. Error bars were obtained from two separate biological

replicates.

Machine learning algorithm optimization for predicting gene function (Chapter 4)

We used the machine learning algorithm, k-nearest nearest neighbors to, in an

unsupervised manner, identify other genes with similar resistance and sensitization patterns

using the scikit-learn Python library (318). However, because of the different selective pressures

exerted by each antibiotic, we cannot use the ratio of reads in the experiment versus the control

that map to each gene as the metric for classification. In addition, we wanted to distinguish

between the two following conditions: 1) A ratio change of 0.1 due to 100 reads in the control

and 10 reads in the experiment and 2) A ratio change of 0.1 due to 1000 reads in the control

and 100 reads in the experiment. If both genes are the same length, option 1 will be much less

relevant than option 2 because 100 reads/gene and 10 reads/gene both correspond to a gene

with a significant fitness defect while a change from 1000 reads/gene to 100 reads/gene is more

likely to be a significant change. Therefore, we converted the ratios into a more appropriate

value. This value, representing fitness under the antibiotic treatment is calculated by multiplying

the converted ratio above by the number of reads mapping to that gene in the treatment

condition, and normalizing to the length of the gene. Then, genes were ordered from smallest to

largest "fitness". To place all the samples on the same scale, the gene with the smallest "fitness"

was given a value of 0, and the gene with the highest "fitness" was given a value of 1. All other

genes were placed in order between these values, in increments that increase by 1/(total

number of genes). This final value, which we call the "normalized fitness value", is the value we

use for the machine learning analysis. Finally, essential genes were removed from the dataset,

and the K-nearest neighbors algorithm was further optimized by adjusting the Minkowski distance metric to output the genes with the most similar resistance/sensitization pattern to the test gene. We identified the five genes (the five nearest neighbors) having the most similar pattern of "normalized fitness values".

Identification of upregulated genes (Chapter 5)

To identify genes upregulated by transposon insertion, the untreated and antibiotic-treated data is separated by type of transposon construct. Then, we separate the data by reads mapping to the plus strand and the minus strand of the genome. This lets us distinguish between the two different directions the outward-facing promoter may be facing. We then map this data to 270bp sections of the genome or windows. Each treated dataset from each strand is normalized to the control using simulation based resampling. Then, we identify the mean and standard deviation of reads that map to each TA site for each strand in the control. Then we look for TA sites in the experiment that are 4-6 standard deviations from the mean of the control. These are potential hits. If there are at least two potential hits in a window or a few windows in a row with hits, we investigate whether there are any genes nearby (within a few kb) that could be upregulated by these transposon insertions. Potential upregulated genes are then investigated to make sure that no reads map to them (if upregulation increases fitness, we would expect that inactivation would decrease the fitness) by mapping the reads/TA site as bar charts in Excel. This chart also allows us to check whether the majority of TA sites in the promoter region of the gene of interest contain transposons with promoters facing the same direction as the gene of interest. For scripts, see Appendix H.

Hierarchical Clustering (Chapter 5)

Hierarchical clustering was performed using the statistical programming language, R, and visualized using the gplots library. The spearman method was used to measure correlation and complete-linkage clustering to measure distance. We used the unique set of genes which had the largest 99th percentile of the data and smallest changes 1st percentile compared to the control to decrease the amount of noise in the data, and we used the "fitness value" described above for each gene to perform the comparison.

K-Nearest Neighbors Algorithm for predicting mechanism of action (Chapter 5)

The sci-kit learn Python library was used for this analysis which came along with the anaconda distribution of the Python language. I wrote Python scripts that convert the raw read count data into fitness values for each antibiotic. This data is then fed into the K-nearest neighbors algorithm in an unsupervised way (where I identify each antibiotic as a unique class) or in a supervised way (where I classify the antibiotics by their mechanism of action) as a training set. We use the minkowski distance metric, a step size = 0.2, and each antibiotic was weighted uniformly. Two neighbors were used to predict the mechanism of an unknown antibiotic. With a larger training data set, this value could be increased which could result in more accurate predictions.

B. Strains and Plasmids

Strain	Genotype/Phenotype	Source (if other than here)
RN4220	S. aureus subsp. aureus NCTC8325 MSSA; r− m+; partial agr defect	Kreiswirth et al. Nature. 1983
HG003	S. aureus subsp. aureus NCTC8325 MSSA ; Φ11 Φ12 Φ113, r+ m+; agr+	Herbert et al. Infect Immun. 2010
TM51	RN4220 attB::Orf5 (pTM378)	Wang et al. Nat Chem Biol. 2011
TM53	RN4220 attB::Orf5 (pTM381)	Wang et al. Nat Chem Biol. 2011
TM174	RN4220 (pOrf5 Tnp+) KmR	
TM175	RN4220 (pOrf5 Tnp-) KmR	
TM176	HG003 (pOrf5 Tnp+) KmR	
TM177	HG003 (pOrf5 Tnp-) KmR	
TM222	HG003 attLint::FRT attR::FRT	
TM226	HG003 Φ11::FRT	
TM231	TM226 (pOrf5 Tnp+)	
TM232	TM226 (pOrf5 Tnp-)	
ΔlyrA	HG003 lyrA::KmR	
ΔmprF	HG003 mprF::KmR	
MW2	S. aureus subsp. aureus Community acquired MRSA: ATCC BAA-1707 ermS, kanS	Baba et al. Lancet. 2002
USA300	S. aureus subsp. aureus Community acquired MRSA strain USA300_TCH1516: ATCC BAA-1717 ermR, kanR, catS	Highlander et al. BMC Microbiol. 2007
TM258	MW2 ΔhsdR - restriction system knocked out, parent strain for making transposon libraries	
TM262	TM258 with pWV01-Tnpase+-ORF5	

TM263	TM258 with pWV01-Tnpase damaged-ORF5	
TM283	USA300 ΔpUSA300HOUMR ermS, kanS, catS, parent strain for making transposon libraries	
TM284	TM283 with pWV01-Tnpase+-ORF5	
TM285	TM283 with pWV01-Tnpase damaged-ORF5	
MW2 Δpbp3	MW2 unmarked pbp3 deletion strain	Memmi et al. Antimicrob Agents Chemother. 2008
MW2 Δpbp4	MW2 unmarked pbp4 deletion strain	Memmi et al. Antimicrob Agents Chemother. 2008
tn::pbp3	Nebraska Transposon library mutant NE420 transduced into TM283	transduced from strain from Fey et al. MBio. 2013
tn::pbp4	Nebraska Transposon library mutant NE679 transduced into TM283	transduced from strain from Fey et al. MBio. 2013
SHM161	HG003 ΔlyrA pTM63-lyrA-cmyc (integrated into genome), camR	
tn::lytD	Nebraska Transposon library mutant NE1909 transduced into TM283	transduced from strain from Fey et al. MBio. 2013
USA300-pLOW-lytD	pLOW-lytD electroporated into TM283	
USA300 ΔlyrA	ΔlyrA (kanR marked) transduced into TM283	
ΔlyrA-pLOW-lytD	pLOW-lytD electroporated into USA300 ΔlyrA	
JE2	WT USA300_FPR3757, parent background fro Nebraska transposon mutant library, ATCC BAA-1556	Fey et al. MBio. 2013
tn::ndh	Nebraska Transposon library mutant NE1884 transduced into JE2	transduced from strain from Fey et al. MBio. 2013
tn::qoxB	Nebraska Transposon library mutant NE732 transduced into JE2	transduced from strain from Fey et al. MBio. 2013
tn::qoxA	Nebraska Transposon library mutant NE92 transduced into JE2	transduced from strain from Fey et al. MBio. 2013
tn::fmtA	Nebraska Transposon library mutant NE1022 transduced into JE2	transduced from strain from Fey et al. MBio. 2013
Newman	MSSA strain Newman, ATCC 25904	Hawiger et al. J Lab Clin Med. 1970
Newman ΔgraR	unmarked graR deletion in Newman	Santa Maria et al. PNAS. 2014
Newman ΔmprF	unmarked mprF deletion in Newman	transduced from strain from Ting Pang, Bernhardt Lab
Newman ΔdltA	unmarked dltA deletion in Newman	Santa Maria et al. PNAS. 2014
tn::1025	Nebraska Transposon library mutant NE1044 transduced into JE2	transduced from strain from Fey et al. MBio. 2013

tn::1050	Nebraska Transposon library mutant NE1420 transduced into JE2	transduced from strain from Fey et al. MBio. 2013
tn::02149	Nebraska Transposon library mutant NE1894 transduced into HG003	transduced from strain from Fey et al. MBio. 2013
tn::00969	Nebraska Transposon library mutant NE50 transduced into HG003	transduced from strain from Fey et al. MBio. 2013
tn::uppP	Nebraska Transposon library mutant NE480 transduced into HG003	transduced from strain from Fey et al. MBio. 2013
HG003 pLOW-02149	HG003 expressing SAOUHSC_02149 from pLOW plasmid	
HG003 pLOW-00969	HG003 expressing SAOUHSC_00969 from pLOW plasmid	
HG003 pLOW-02164	HG003 expressing SAOUHSC_02164 from pLOW plasmid	
HG003 pLOW-uppP	HG003 expressing SAOUHSC_00691 (uppP) from pLOW plasmid	

Plasmids	Genotype/Phenotype	Source (if other than here)
pKFC	temperature-sensitive shuttle vector; Ampr in E. coli, Cmr in S aureus	Kate and Sugai. J Microbiol Methods. 2011
pUC19	pMB1 ori lacZ' ApR Clontech pTM402 pT181 repC- cop	Clontech
pTM402	pT181 repC- cop 623 ori sso; 1-kb DNA Φ 11 fragment; mini-Tnp cassette with ermB of Tn551 and outward facing Ppen promoter; Em R	Wang et al. Nat Chem Biol. 2011
pTM378	pWV01ts ori aphA-3 Gram+ RBS HMAR1 C9 transposase ; KmR	Wang et al. Nat Chem Biol. 2011
pTM381	pWV01ts ori aphA-3 Gram+ RBS ΔHMAR1 C9 truncated transposase ; KmR	Wang et al. Nat Chem Biol. 2011
pCP20	Vector containing the S. cervesiae FLP recombinase	Cerepanov and Wackernagel. Gene. 1995
pMS182	E. coli/S. aureus shuttle vector Cmr pLI50 with Ppen GFP-mut2	Swoboda et al. ACS Chem Biol. 2009
pORF5 Tnp+	pTM378 with the cI-like repressor ORF5 from Φ11	
pORF5 Tnp-	pTM381 with the cI-like repressor ORF5 from Φ11	
pTM239 (Blunt)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gGTAa) - EmR	
pTM240 (Ppen)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gAATa) and outward facing Ppen promoter- EmR	

pTM241 (Pcap)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gTGGa) for outward facing Pcap promoter- EmR	
pTM242 (Ptuf)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gGATa) for outward facing Ptuf promoter- EmR	
pTM243 (Perm)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gGCAa) for outward facing Perm promoter- EmR	
pTM244 (Pdual)	pTM402 modified with NotI-P7 annealing site-ITR2::MmeI-NotI with DNA barcode (gATTa) for outward facing PDual promoter- EmR	
pTM204*attLint*	pKFC vector with 1-kb DNA homology regions flanking attLint on each side of FRT recombination sequence	
pTM204*attR*	pKFC vector with 1-kb DNA homology regions flanking attR on each side of FRT recombination sequence	
pTM195	pLI50- Ppen-FLP recombinase of Saccharomyces cervesiae	
(pKFC-lyrA)	pKFC vector with 1-kb DNA homology regions flanking lyrA on each side of the kanamycin resistance cassette	Santa Maria et al. PNAS. 2014
(pKFC-mprF)	pKFC vector with 1-kb DNA homology regions flanking mprF on each side of the kanamycin resistance cassette	
pKFC-hsdR (MW016)	pKFC vector for in fame knockout of hsdR using Clontech assembly	
pTM283	pKFC vector with in frame deletion of ermR allele	
pTP63	vector for single copy integration of gene/construct into genome	Ting Pang, Bernhardt lab
pTP44	same as pRAB14, encodes integrase for integration of pTP63 into *S. aureus* genome	Ting Pang, Bernhardt lab
pLOW-lytD	USA300HOU_1765 cloned into pLOW plasmid (constitutive upregulation), ermR	
pLOW-02149	SAOUHSC_02149 cloned into pLOW plasmid (constitutive upregulation), ermR	
pLOW-02164	SAOUHSC_02164 cloned into pLOW plasmid (constitutive upregulation), ermR	
pLOW-00969	SAOUHSC_00969 cloned into pLOW plasmid (constitutive upregulation), ermR	

pLOW-uppP | uppP (SAOUHSC_00691) cloned into pLOW plasmid (constitutive upregulation), ermR

C. Detailed protocol for the preparation of transposon library DNA for NGS

Unless otherwise described, reactions are mixed and incubated in 1.5 mL Eppendorf LoBind tubes. DNA was stored at -20°C overnight, and at -80°C long term. At least 10 μg of high molecular weight genomic DNA was purified. 10 μg of genomic DNA was digested with 50-100 U NotI in a 600 μl reaction in NEB Buffer #3 supplemented with BSA. The reaction was vortexed gently, mixed by inversion, spun down, and incubated at 37°C for seven hours (mixing and spinning down once halfway through). NotI was inactivated at 70°C for 20 minutes, and cooled to room temperature (5 minutes).

Next, the transposon-plasmid junctions were removed through a size-selective precipitation. The 4x Precipitation buffer contains 32% PEG8000, 2.2M NaCl (autoclaved), 40mM Na2PO4 or K2PO4 pH 7.5 (autoclaved), and brought to volume with autoclaved ddH2O. 200μl of 4x precipitation buffer was added to the 600μl digest reaction and vortexed and inverted (not pipetted) to mix. This reaction was incubated in an ice water bath in a 4°C cold room for 12-16 hours. After incubation, the reaction was spun down in a tabletop centrifuge at maximum speed for 20 minutes at 4°C. Then, the supernatant was removed, and the precipitated DNA was washed once with cold 1x Precipitation buffer, spinning down the DNA at maximum speed for 10 minutes at 4°C. To further purify the DNA, pellets were washed twice with room temperature 70% ethanol, spinning down the DNA at maximum speed for 5 minutes at room temperature between washes. The final DNA pellet was dried and resuspended in 50-100μl standard elution buffer (10mM Tris, pH 8.5) by pipetting.

A PCR check was performed on one sample (undigested and NotI digested) to confirm that the DNA was sufficiently digested using primers that anneal to the transposon on either side of a NotI site. As a control, other primers were also used that anneal to a genomic fragment inside the tarK gene. Seven reactions were set up for each primer set for both undigested and

206

digested DNA. The 25µl PCR reaction consisted of 2.5µl 10x KOD Hot Start Buffer, 5 µl 5x CES

mix, 1.5µl 25mM MgSO4, 2.5 µl 2mM dNTPs, 0.25 µl of each primer (experimental or control),

0.25 µl KOD Hot Start Polymerase, 25 ng of DNA (digested or undigested), and autoclaved

ddH2O to bring the reaction to 25 µl. Reactions were incubated in a thermocycler at 95°C for 2

minutes; with 30 cycles of 95°C for 20 seconds, 55°C for 20 seconds, and 70°C for 30 seconds;

and 12°C for storage. CES buffer consists of 2.7M betaine, 6.7mM DTT, 6.7% DMSO, and

55µg/ml BSA. Beginning at the end of 9 cycles, a tube from each category of reaction

(undigested with experimental primers, digested with experimental primers, undigested with

control primers, digested with control primers) was removed from the thermocycler every three

cycles and quenched with DNA gel loading buffer supplemented with 0.1% SDS. Samples were

stored at -20°C until all cycles were completed. The final tubes were removed at the end of the

27th cycle. These reactions were run on a 1% agarose gel in TAE buffer. In a sample deemed

sufficiently digested, a six cycle product detection threshold difference between the undigested

and digested samples when using the experimental primers (785 bp) in comparison to the

control primers (1298 bp) was required.

Once we confirmed that the DNA had been sufficiently digested, the biotinylated

adaptors were ligated. 50 µM of annealed adaptors were prepared by mixing 15 µl of 100 µM

TM214 and 15 µl of 100 µM TM215 with 1.5 µl 1M NaCl. This reaction was incubated in a

thermocycler at 95°C for 5 minutes followed by cooling to 4°C at a rate of 0.1°C/second. Each

150µl ligation reaction consisted of 15 µl of 10x T4 ligase buffer, 8 µg digested DNA in 100 µl of

ddH2O, 3 µl of T4 ligase enzyme, 4.5 µl of annealed adaptors diluted 1:10 to 5 µM in ice cold 1x

T4 ligase buffer, and 27.5 µl autoclaved ddH2O. Reactions were incubated at 16°C overnight.

Another size-selective precipitation was performed to remove un-ligated adaptors from genomic

DNA. 50 µl of the same 4x PEG solution described above was added to the ligation reaction,

and the tube mixed by vortexing, spun down, and incubated at 4°C or in the cold room for 12-16

hours. After incubation, the DNA is washed and dried in the same manner as described above after the NotI digest.

Next, transposon-genome junctions were captured by digesting the DNA with MmeI which digests 20 bp 3' from its asymmetric recognition site near the end of the transposon ITR. The M12 oligonucleotides were first annealed together as described above to create 50 μM double-stranded DNA. MmeI digests take place in 200μl reactions and require at least 5 μg of DNA from the previous step diluted to 50 μl in autoclaved ddH2O. MmeI reactions consist of 20 μl 10x NEB CutSmart Buffer, 0.8 μl 32 mM SAM, 2 μl 50 μM annealed M12 oligos, 50 μl DNA, 2 μl MmeI enzyme, and 125.2 μl autoclaved ddH2O. The reactions were incubated at 37°C for two hours.

MmeI-digested DNA was ligated to streptavidin dynabeads via the biotinylated adaptor. This required three buffers. The 2x BandW Buffer consists of 2 M NaCl, 10 mM Tris, and 1 mM EDTA pH 7.5 with concentrated HCl [9]. LoTE Buffer consists of 3 mM Tris, 0.2 mM EDTA pH 7.5 with concentrated HCl [9]. LoTE+Tween Buffer is the same as LoTE buffer, but supplemented with 0.05% Tween 20. 200 μl of 2x BandW buffer was added to each sample. 32 μl/sample of Dynabeads® M-280 Streptavidin beads was added to a LoBind tube and placed in the magnetic particle collector (MPC). Then, the supernatant was removed, and the beads were washed three times with 1 mL of 1x BandW buffer. Finally, the sample was resuspended in 32 μl/sample of 1x BandW buffer, and 32 μl of beads were added to each diluted MmeI digest sample. These were incubated at room temperature for one hour, resuspending by tapping and inversion every 10-15 minutes. At this point, all biotinylated DNA should be bound to the beads, so the beads were collected using the MPC, washed once with LoTE+Tween and twice with LoTE (no Tween). Finally, the beads were resuspended in 100μl of LoTE, and transferred to a PCR tube. Beads can be stored in LoTE buffer at 4°C overnight.

With one end of the DNA attached to beads, the other Illumina adaptors with index barcodes were ligated to the other end. We used six adaptors (LIB_AdaptT_1_long, LIB_AdaptB_1_long, LIB_AdaptT_2_long, LIB_AdaptB_2_long, LIB_AdaptT_3_long, LIB_AdaptB_3_long, LIB_AdaptT_4_long, LIB_AdaptB_4_long, LIB_AdaptT_5_long, LIB_AdaptB_5_long, LIB_AdaptT_6_long, LIB_AdaptB_6_long) annealed to each other (T to B) as described above. We would like to note that neither the barcodes in these adaptors nor the transposon construct-specific barcodes use an error-correcting barcode sequence. Without error correcting barcodes, it is possible that sequences could be mis-assigned to the wrong sample or wrong transposon construct due to errors in sequencing. However, because we only used sequences with a high quality score in our analysis, we assume that the fraction of mis-assigned sequences is negligible. These double-stranded adaptors were diluted tenfold to 5 µM in ice cold T4 ligase buffer. The ligation mix consisted of 16.4 µl ddH2O and 2 µl 10x T4 ligase buffer per sample. Beads were collected in the MPC and resuspended in 18.4 µl ligation mix. 0.6 µl of the double-stranded adaptor is added to each tube, a different adaptor to each tube. Then 1 µl of T4 ligase is added to each tube. The reaction is mixed by pipetting, and the reactions were incubated for 6-7 hours at 16°C in a thermocycler, mixing by tapping and pipetting every 15 minutes. After incubation, the beads were washed once with 150 µl LoTE+Tween and twice with 150 µl LoTE. Again, beads can be stored in LoTE buffer at 4°C overnight.

The final PCR reaction amplifies the transposon-genome junction off of the beads and adds the adaptor sequences required for Illumina sequencing. Beads were collected in the MPC and resuspended in 50 µl of the final PCR reaction mix. This 50 µl PCR reaction consists of 5 µl 10x KOD buffer, 3 µl 25 mM MgSO4, 5 µl 2 mM dNTPs, 35.5 µl autoclaved ddH2O, 0.5 µl 100µM TM199, 0.5 µl 100 µM LIB-PCR_3, and 0.5 µl KOD Hot Start Polymerase per sample. The reaction was incubated in the thermocycler using the following program: 95°C for 2

minutes; 15- 18 cycles of 95°C for 20 seconds, 60°C for 20 seconds, 72°C for 20 seconds. The

beads were collected, and the supernatant was transferred to new tubes where DNA loading

buffer was added, and the samples were run on a 2% agarose gel in TAE at 130V for 20

minutes. Bands of the expected size (161 bp) were extracted using the Qiagen gel extraction kit

with a few minor modifications. The gel fragments were dissolved in Buffer QG at room

temperature to prevent dissociation of the short strands of DNA. After the final Buffer PE wash,

excess buffer was removed from the column by pipette. The DNA was eluted in 30 µl Buffer EB,

and DNA concentration was determined using the Quant-ITTM PicoGreen kit from Invitrogen.

The final TnSeq sample was diluted to 10 nM in Buffer EB, and six samples with different

barcodes were typically multiplexed together in a single lane. Samples were sequenced on a Hi-

Seq2000 or HiSeq2500 for 100 cycles with 40% ΦX174 spiked in to the sequencing reaction.

D. Essential genes with different promoter constructs

Blunt Gene Locus	Erm Gene Locus	Promoters Gene Locus
SAOUHSC_00001-dnaA	SAOUHSC_00001-dnaA	SAOUHSC_00001-dnaA
SAOUHSC_00005	SAOUHSC_00004-recF	SAOUHSC_00004-recF
SAOUHSC_00006	SAOUHSC_00005	SAOUHSC_00005
SAOUHSC_00009	SAOUHSC_00006	SAOUHSC_00006
SAOUHSC_00015	SAOUHSC_00009	SAOUHSC_00009
SAOUHSC_00017-rplI	SAOUHSC_00015	SAOUHSC_00015
SAOUHSC_00018	SAOUHSC_00018	SAOUHSC_00018
SAOUHSC_00021	SAOUHSC_00021	SAOUHSC_00021
SAOUHSC_00226	SAOUHSC_00226	SAOUHSC_00226
SAOUHSC_00349	SAOUHSC_00349	SAOUHSC_00349
SAOUHSC_00350-rpsR	SAOUHSC_00350-rpsR	SAOUHSC_00350-rpsR
SAOUHSC_00374	SAOUHSC_00375-guaA	SAOUHSC_00374
SAOUHSC_00375-guaA	SAOUHSC_00413	SAOUHSC_00375-guaA
SAOUHSC_00413	SAOUHSC_00442	SAOUHSC_00413
SAOUHSC_00442	SAOUHSC_00444	SAOUHSC_00420
SAOUHSC_00444	SAOUHSC_00451-tmk	SAOUHSC_00442
SAOUHSC_00445-recR	SAOUHSC_00454	SAOUHSC_00451-tmk
SAOUHSC_00451-tmk	SAOUHSC_00461	SAOUHSC_00461
SAOUHSC_00454	SAOUHSC_00471-glmU	SAOUHSC_00471-glmU
SAOUHSC_00461	SAOUHSC_00472	SAOUHSC_00482
SAOUHSC_00471-glmU	SAOUHSC_00482	SAOUHSC_00484
SAOUHSC_00472	SAOUHSC_00484	SAOUHSC_00485
SAOUHSC_00481	SAOUHSC_00485	SAOUHSC_00490
SAOUHSC_00482	SAOUHSC_00490	SAOUHSC_00491
SAOUHSC_00484	SAOUHSC_00491	SAOUHSC_00493-lysS
SAOUHSC_00485	SAOUHSC_00493-lysS	SAOUHSC_00511-cysS
SAOUHSC_00486-ftsH	SAOUHSC_00511-cysS	SAOUHSC_00518-rplK
SAOUHSC_00490	SAOUHSC_00518-rplK	SAOUHSC_00519-rplA
SAOUHSC_00491	SAOUHSC_00519-rplA	SAOUHSC_00520-rplJ
SAOUHSC_00493-lysS	SAOUHSC_00520-rplJ	SAOUHSC_00521-rplL
SAOUHSC_00510	SAOUHSC_00521-rplL	SAOUHSC_00524-rpoB
SAOUHSC_00511-cysS	SAOUHSC_00524-rpoB	SAOUHSC_00525
SAOUHSC_00518-rplK	SAOUHSC_00525	SAOUHSC_00527-rpsL
SAOUHSC_00519-rplA	SAOUHSC_00526	SAOUHSC_00528
SAOUHSC_00520-rplJ	SAOUHSC_00527-rpsL	SAOUHSC_00529-fusA
SAOUHSC_00521-rplL	SAOUHSC_00528	SAOUHSC_00530-tuf
SAOUHSC_00524-rpoB	SAOUHSC_00529-fusA	SAOUHSC_00574-eutD

SAOUHSC_00525   SAOUHSC_00530-tuf   SAOUHSC_00575
SAOUHSC_00527-rpsL   SAOUHSC_00575   SAOUHSC_00578
SAOUHSC_00528   SAOUHSC_00578   SAOUHSC_00579
SAOUHSC_00529-fusA   SAOUHSC_00579   SAOUHSC_00611-argS
SAOUHSC_00530-tuf   SAOUHSC_00611-argS   SAOUHSC_00641-tarB
SAOUHSC_00574-eutD   SAOUHSC_00641-tarB   SAOUHSC_00643-tarH
SAOUHSC_00575   SAOUHSC_00643-tarH   SAOUHSC_00645-tarD
SAOUHSC_00578   SAOUHSC_00645-tarD   SAOUHSC_00742
SAOUHSC_00579   SAOUHSC_00742   SAOUHSC_00743-nrdF
SAOUHSC_00610   SAOUHSC_00743-nrdF   SAOUHSC_00752-murB
SAOUHSC_00611-argS   SAOUHSC_00752-murB   SAOUHSC_00771
SAOUHSC_00641-tarB   SAOUHSC_00771   SAOUHSC_00778
SAOUHSC_00643-tarH   SAOUHSC_00778   SAOUHSC_00785
SAOUHSC_00645-tarD   SAOUHSC_00785   SAOUHSC_00788
SAOUHSC_00742   SAOUHSC_00788   SAOUHSC_00795
SAOUHSC_00743-nrdF   SAOUHSC_00789   SAOUHSC_00796-pgk
SAOUHSC_00752-murB   SAOUHSC_00795   SAOUHSC_00797-tpiA
SAOUHSC_00769-secA   SAOUHSC_00796-pgk   SAOUHSC_00798
SAOUHSC_00771   SAOUHSC_00797-tpiA   SAOUHSC_00799-eno
SAOUHSC_00778   SAOUHSC_00798   SAOUHSC_00804-smpB
SAOUHSC_00785   SAOUHSC_00799-eno   SAOUHSC_00832
SAOUHSC_00788   SAOUHSC_00804-smpB   SAOUHSC_00848
SAOUHSC_00795   SAOUHSC_00832   SAOUHSC_00849
SAOUHSC_00796-pgk   SAOUHSC_00848   SAOUHSC_00850
SAOUHSC_00797-tpiA   SAOUHSC_00849   SAOUHSC_00851
SAOUHSC_00798   SAOUHSC_00850   SAOUHSC_00868
SAOUHSC_00799-eno   SAOUHSC_00851   SAOUHSC_00869-dltA
SAOUHSC_00804-smpB   SAOUHSC_00868   SAOUHSC_00870-dltB
SAOUHSC_00832   SAOUHSC_00869-dltA   SAOUHSC_00871-dltC
SAOUHSC_00848   SAOUHSC_00870-dltB   SAOUHSC_00872-dltD
SAOUHSC_00849   SAOUHSC_00871-dltC   SAOUHSC_00884
SAOUHSC_00850   SAOUHSC_00872-dltD   SAOUHSC_00885
SAOUHSC_00851   SAOUHSC_00884   SAOUHSC_00886
SAOUHSC_00868   SAOUHSC_00885   SAOUHSC_00887
SAOUHSC_00869-dltA   SAOUHSC_00886   SAOUHSC_00888
SAOUHSC_00870-dltB   SAOUHSC_00887   SAOUHSC_00889
SAOUHSC_00871-dltC   SAOUHSC_00888   SAOUHSC_00900-pgi
SAOUHSC_00872-dltD   SAOUHSC_00889   SAOUHSC_00903
SAOUHSC_00884   SAOUHSC_00900-pgi   SAOUHSC_00904
SAOUHSC_00885   SAOUHSC_00903   SAOUHSC_00905
SAOUHSC_00886   SAOUHSC_00904   SAOUHSC_00921
SAOUHSC_00887   SAOUHSC_00905   SAOUHSC_00933

SAOUHSC_00888
SAOUHSC_00889
SAOUHSC_00900-pgi
SAOUHSC_00903
SAOUHSC_00904
SAOUHSC_00905
SAOUHSC_00921
SAOUHSC_00933
SAOUHSC_00943-ppnK
SAOUHSC_00947
SAOUHSC_00953-ypfP
SAOUHSC_00980-menA
SAOUHSC_00981
SAOUHSC_00982-menF
SAOUHSC_00983-menD
SAOUHSC_00985
SAOUHSC_01000
SAOUHSC_01002
SAOUHSC_01035
SAOUHSC_01036
SAOUHSC_01040
SAOUHSC_01041
SAOUHSC_01042
SAOUHSC_01043
SAOUHSC_01063-ftsW
SAOUHSC_01075-coaD
SAOUHSC_01077
SAOUHSC_01078
SAOUHSC_01092-pheS
SAOUHSC_01093-pheT
SAOUHSC_01142
SAOUHSC_01143-mraW
SAOUHSC_01144-ftsL
SAOUHSC_01145
SAOUHSC_01146-mraY
SAOUHSC_01147-murD
SAOUHSC_01148-div1B
SAOUHSC_01149
SAOUHSC_01150-ftsZ
SAOUHSC_01158
SAOUHSC_01159-ileS
SAOUHSC_01163

SAOUHSC_00921
SAOUHSC_00933
SAOUHSC_00943-ppnK
SAOUHSC_00947
SAOUHSC_00953-ypfP
SAOUHSC_00980-menA
SAOUHSC_00981
SAOUHSC_00982-menF
SAOUHSC_00983-menD
SAOUHSC_00984
SAOUHSC_00985
SAOUHSC_01000
SAOUHSC_01035
SAOUHSC_01036
SAOUHSC_01040
SAOUHSC_01041
SAOUHSC_01042
SAOUHSC_01043
SAOUHSC_01063-ftsW
SAOUHSC_01075-coaD
SAOUHSC_01078
SAOUHSC_01093-pheT
SAOUHSC_01144-ftsL
SAOUHSC_01145
SAOUHSC_01146-mraY
SAOUHSC_01147-murD
SAOUHSC_01148-div1B
SAOUHSC_01149
SAOUHSC_01150-ftsZ
SAOUHSC_01159-ileS
SAOUHSC_01163
SAOUHSC_01166-pyrB
SAOUHSC_01168-pyrC
SAOUHSC_01169
SAOUHSC_01170-carB
SAOUHSC_01171
SAOUHSC_01172-pyrE
SAOUHSC_01178
SAOUHSC_01179
SAOUHSC_01183
SAOUHSC_01188
SAOUHSC_01189

SAOUHSC_00943-ppnK
SAOUHSC_00953-ypfP
SAOUHSC_00980-menA
SAOUHSC_00982-menF
SAOUHSC_00983-menD
SAOUHSC_00985
SAOUHSC_01035
SAOUHSC_01036
SAOUHSC_01040
SAOUHSC_01041
SAOUHSC_01042
SAOUHSC_01043
SAOUHSC_01063-ftsW
SAOUHSC_01075-coaD
SAOUHSC_01093-pheT
SAOUHSC_01144-ftsL
SAOUHSC_01145
SAOUHSC_01146-mraY
SAOUHSC_01147-murD
SAOUHSC_01148-div1B
SAOUHSC_01149
SAOUHSC_01150-ftsZ
SAOUHSC_01159-ileS
SAOUHSC_01163
SAOUHSC_01166-pyrB
SAOUHSC_01168-pyrC
SAOUHSC_01169
SAOUHSC_01170-carB
SAOUHSC_01172-pyrE
SAOUHSC_01179
SAOUHSC_01183
SAOUHSC_01189
SAOUHSC_01190
SAOUHSC_01194
SAOUHSC_01197
SAOUHSC_01198
SAOUHSC_01199
SAOUHSC_01201-acpP
SAOUHSC_01205
SAOUHSC_01207
SAOUHSC_01210-trmD
SAOUHSC_01211-rplS

SAOUHSC_01166-pyrB	SAOUHSC_01190	SAOUHSC_01216-sucC
SAOUHSC_01168-pyrC	SAOUHSC_01194	SAOUHSC_01222
SAOUHSC_01169	SAOUHSC_01197	SAOUHSC_01228-codY
SAOUHSC_01170-carB	SAOUHSC_01198	SAOUHSC_01233
SAOUHSC_01171	SAOUHSC_01199	SAOUHSC_01234-tsf
SAOUHSC_01172-pyrE	SAOUHSC_01201-acpP	SAOUHSC_01235-pyrH
SAOUHSC_01178	SAOUHSC_01205	SAOUHSC_01236-frr
SAOUHSC_01179	SAOUHSC_01207	SAOUHSC_01238
SAOUHSC_01183	SAOUHSC_01210-trmD	SAOUHSC_01240
SAOUHSC_01188	SAOUHSC_01211-rplS	SAOUHSC_01241-polC
SAOUHSC_01189	SAOUHSC_01216-sucC	SAOUHSC_01243-nusA
SAOUHSC_01190	SAOUHSC_01218	SAOUHSC_01244
SAOUHSC_01194	SAOUHSC_01222	SAOUHSC_01245
SAOUHSC_01197	SAOUHSC_01228-codY	SAOUHSC_01246-infB
SAOUHSC_01198	SAOUHSC_01234-tsf	SAOUHSC_01249
SAOUHSC_01199	SAOUHSC_01235-pyrH	SAOUHSC_01250-rpsO
SAOUHSC_01200	SAOUHSC_01236-frr	SAOUHSC_01252
SAOUHSC_01201-acpP	SAOUHSC_01238	SAOUHSC_01260
SAOUHSC_01205	SAOUHSC_01240	SAOUHSC_01262-recA
SAOUHSC_01207	SAOUHSC_01241-polC	SAOUHSC_01287
SAOUHSC_01209-rimM	SAOUHSC_01243-nusA	SAOUHSC_01337
SAOUHSC_01210-trmD	SAOUHSC_01244	SAOUHSC_01351
SAOUHSC_01211-rplS	SAOUHSC_01245	SAOUHSC_01352
SAOUHSC_01216-sucC	SAOUHSC_01246-infB	SAOUHSC_01373
SAOUHSC_01218	SAOUHSC_01249	SAOUHSC_01374
SAOUHSC_01222	SAOUHSC_01250-rpsO	SAOUHSC_01424-murG
SAOUHSC_01228-codY	SAOUHSC_01252	SAOUHSC_01434
SAOUHSC_01230	SAOUHSC_01260	SAOUHSC_01467
SAOUHSC_01234-tsf	SAOUHSC_01262-recA	SAOUHSC_01470
SAOUHSC_01235-pyrH	SAOUHSC_01287	SAOUHSC_01473
SAOUHSC_01236-frr	SAOUHSC_01337	SAOUHSC_01474
SAOUHSC_01238	SAOUHSC_01351	SAOUHSC_01481
SAOUHSC_01239	SAOUHSC_01352	SAOUHSC_01482-aroB
SAOUHSC_01240	SAOUHSC_01373	SAOUHSC_01483-aroF
SAOUHSC_01241-polC	SAOUHSC_01374	SAOUHSC_01487-ubiE
SAOUHSC_01243-nusA	SAOUHSC_01424-murG	SAOUHSC_01491-gpsA
SAOUHSC_01244	SAOUHSC_01434	SAOUHSC_01492-engA
SAOUHSC_01245	SAOUHSC_01467	SAOUHSC_01496-cmk
SAOUHSC_01246-infB	SAOUHSC_01470	SAOUHSC_01501
SAOUHSC_01249	SAOUHSC_01473	SAOUHSC_01591
SAOUHSC_01250-rpsO	SAOUHSC_01474	SAOUHSC_01592-fur
SAOUHSC_01252	SAOUHSC_01481	SAOUHSC_01598

SAOUHSC_01259

SAOUHSC_01260

SAOUHSC_01262-recA

SAOUHSC_01265

SAOUHSC_01267

SAOUHSC_01287

SAOUHSC_01337

SAOUHSC_01351

SAOUHSC_01352

SAOUHSC_01373

SAOUHSC_01374

SAOUHSC_01424-murG

SAOUHSC_01434

SAOUHSC_01467

SAOUHSC_01470

SAOUHSC_01473

SAOUHSC_01474

SAOUHSC_01481

SAOUHSC_01482-aroB

SAOUHSC_01483-aroF

SAOUHSC_01487-ubiE

SAOUHSC_01488

SAOUHSC_01491-gpsA

SAOUHSC_01492-engA

SAOUHSC_01496-cmk

SAOUHSC_01501

SAOUHSC_01591

SAOUHSC_01592-fur

SAOUHSC_01598

SAOUHSC_01599

SAOUHSC_01612-bkdA2

SAOUHSC_01613-bkdA1

SAOUHSC_01614

SAOUHSC_01619

SAOUHSC_01623

SAOUHSC_01624

SAOUHSC_01635

SAOUHSC_01659

SAOUHSC_01661

SAOUHSC_01662

SAOUHSC_01663

SAOUHSC_01666

SAOUHSC_01482-aroB

SAOUHSC_01483-aroF

SAOUHSC_01487-ubiE

SAOUHSC_01492-engA

SAOUHSC_01501

SAOUHSC_01550

SAOUHSC_01553

SAOUHSC_01555

SAOUHSC_01556

SAOUHSC_01567

SAOUHSC_01575

SAOUHSC_01591

SAOUHSC_01592-fur

SAOUHSC_01598

SAOUHSC_01612-bkdA2

SAOUHSC_01613-bkdA1

SAOUHSC_01623

SAOUHSC_01624

SAOUHSC_01659

SAOUHSC_01661

SAOUHSC_01662

SAOUHSC_01663

SAOUHSC_01666

SAOUHSC_01668-era

SAOUHSC_01672

SAOUHSC_01690-holA

SAOUHSC_01697

SAOUHSC_01698

SAOUHSC_01700

SAOUHSC_01701

SAOUHSC_01702

SAOUHSC_01720

SAOUHSC_01721

SAOUHSC_01722-alaS

SAOUHSC_01727

SAOUHSC_01738-hisS

SAOUHSC_01746

SAOUHSC_01753-obgE

SAOUHSC_01756

SAOUHSC_01757-rplU

SAOUHSC_01766

SAOUHSC_01767-valS

SAOUHSC_01612-bkdA2

SAOUHSC_01613-bkdA1

SAOUHSC_01619

SAOUHSC_01621-nusB

SAOUHSC_01623

SAOUHSC_01624

SAOUHSC_01659

SAOUHSC_01661

SAOUHSC_01662

SAOUHSC_01663

SAOUHSC_01666

SAOUHSC_01668-era

SAOUHSC_01672

SAOUHSC_01690-holA

SAOUHSC_01697

SAOUHSC_01698

SAOUHSC_01700

SAOUHSC_01701

SAOUHSC_01720

SAOUHSC_01721

SAOUHSC_01722-alaS

SAOUHSC_01727

SAOUHSC_01738-hisS

SAOUHSC_01746

SAOUHSC_01753-obgE

SAOUHSC_01756

SAOUHSC_01757-rplU

SAOUHSC_01766

SAOUHSC_01767-valS

SAOUHSC_01772

SAOUHSC_01773

SAOUHSC_01774-hemC

SAOUHSC_01776-hemA

SAOUHSC_01782

SAOUHSC_01784-rplT

SAOUHSC_01785-rpmI

SAOUHSC_01786-infC

SAOUHSC_01787

SAOUHSC_01791

SAOUHSC_01792

SAOUHSC_01807

SAOUHSC_01809

SAOUHSC_01668-era SAOUHSC_01772 SAOUHSC_01826

SAOUHSC_01670 SAOUHSC_01773 SAOUHSC_01839

SAOUHSC_01672 SAOUHSC_01774-hemC SAOUHSC_01865-trmB

SAOUHSC_01690-holA SAOUHSC_01776-hemA SAOUHSC_01858

SAOUHSC_01697 SAOUHSC_01784-rplT SAOUHSC_01866

SAOUHSC_01698 SAOUHSC_01785-rpmI SAOUHSC_01871

SAOUHSC_01699-aroE SAOUHSC_01786-infC SAOUHSC_01875-leuS

SAOUHSC_01700 SAOUHSC_01787 SAOUHSC_01915-menC

SAOUHSC_01701 SAOUHSC_01791 SAOUHSC_01916

SAOUHSC_01720 SAOUHSC_01792 SAOUHSC_01961-hemH

SAOUHSC_01721 SAOUHSC_01807 SAOUHSC_01962-hemE

SAOUHSC_01722-alaS SAOUHSC_01809 SAOUHSC_01973

SAOUHSC_01727 SAOUHSC_01811 SAOUHSC_02106

SAOUHSC_01738-hisS SAOUHSC_01812 SAOUHSC_02107

SAOUHSC_01741 SAOUHSC_01826 SAOUHSC_02114

SAOUHSC_01746 SAOUHSC_01839 SAOUHSC_02116-gatB

SAOUHSC_01750-ruvB SAOUHSC_01865-trmB SAOUHSC_02117-gatA

SAOUHSC_01753-obgE SAOUHSC_01866 SAOUHSC_02118-gatC

SAOUHSC_01756 SAOUHSC_01871 SAOUHSC_02122

SAOUHSC_01757-rplU SAOUHSC_01875-leuS SAOUHSC_02123

SAOUHSC_01766 SAOUHSC_01909 SAOUHSC_02133

SAOUHSC_01767-valS SAOUHSC_01915-menC SAOUHSC_02140

SAOUHSC_01772 SAOUHSC_01916 SAOUHSC_02143

SAOUHSC_01773 SAOUHSC_01961-hemH SAOUHSC_02151

SAOUHSC_01774-hemC SAOUHSC_01962-hemE SAOUHSC_02152

SAOUHSC_01776-hemA SAOUHSC_01973 SAOUHSC_02228

SAOUHSC_01778-clpX SAOUHSC_02106 SAOUHSC_02277-gcp

SAOUHSC_01782 SAOUHSC_02107 SAOUHSC_02279

SAOUHSC_01783 SAOUHSC_02114 SAOUHSC_02280

SAOUHSC_01784-rplT SAOUHSC_02116-gatB SAOUHSC_02318-ddl

SAOUHSC_01785-rpmI SAOUHSC_02117-gatA SAOUHSC_02336-fabZ

SAOUHSC_01786-infC SAOUHSC_02118-gatC SAOUHSC_02357

SAOUHSC_01787 SAOUHSC_02122 SAOUHSC_02359-prfA

SAOUHSC_01789 SAOUHSC_02123 SAOUHSC_02369

SAOUHSC_01791 SAOUHSC_02133 SAOUHSC_02400

SAOUHSC_01792 SAOUHSC_02140 SAOUHSC_02409

SAOUHSC_01793-nrdR SAOUHSC_02143 SAOUHSC_02480-truA

SAOUHSC_01807 SAOUHSC_02151 SAOUHSC_02486

SAOUHSC_01809 SAOUHSC_02152 SAOUHSC_02487-rpsM

SAOUHSC_01811 SAOUHSC_02219 SAOUHSC_02488-rpmJ

SAOUHSC_01828 SAOUHSC_02220 SAOUHSC_02489-infA

SAOUHSC_01839 SAOUHSC_02227 SAOUHSC_02490-adk

SAOUHSC_01865-trmB
SAOUHSC_01858
SAOUHSC_01866
SAOUHSC_01871
SAOUHSC_01875-leuS
SAOUHSC_01915-menC
SAOUHSC_01916
SAOUHSC_01961-hemH
SAOUHSC_01962-hemE
SAOUHSC_01973
SAOUHSC_02106
SAOUHSC_02107
SAOUHSC_02114
SAOUHSC_02116-gatB
SAOUHSC_02117-gatA
SAOUHSC_02118-gatC
SAOUHSC_02122
SAOUHSC_02123
SAOUHSC_02133
SAOUHSC_02140
SAOUHSC_02143
SAOUHSC_02151
SAOUHSC_02152
SAOUHSC_02255-groES
SAOUHSC_02277-gcp
SAOUHSC_02278
SAOUHSC_02279
SAOUHSC_02280
SAOUHSC_02318-ddl
SAOUHSC_02336-fabZ
SAOUHSC_02357
SAOUHSC_02359-prfA
SAOUHSC_02369
SAOUHSC_02400
SAOUHSC_02409
SAOUHSC_02480-truA
SAOUHSC_02483-cbiO
SAOUHSC_02486
SAOUHSC_02487-rpsM
SAOUHSC_02488-rpmJ
SAOUHSC_02489-infA
SAOUHSC_02490-adk

SAOUHSC_02228
SAOUHSC_02233
SAOUHSC_02255-groES
SAOUHSC_02277-gcp
SAOUHSC_02279
SAOUHSC_02280
SAOUHSC_02318-ddl
SAOUHSC_02336-fabZ
SAOUHSC_02357
SAOUHSC_02359-prfA
SAOUHSC_02369
SAOUHSC_02400
SAOUHSC_02409
SAOUHSC_02480-truA
SAOUHSC_02486
SAOUHSC_02487-rpsM
SAOUHSC_02488-rpmJ
SAOUHSC_02489-infA
SAOUHSC_02490-adk
SAOUHSC_02491-secY
SAOUHSC_02492-rplO
SAOUHSC_02493-rpmD
SAOUHSC_02494-rpsE
SAOUHSC_02495-rplR
SAOUHSC_02496-rplF
SAOUHSC_02498-rpsH
SAOUHSC_02499-rpsN
SAOUHSC_02500-rplE
SAOUHSC_02501-rplX
SAOUHSC_02502-rplN
SAOUHSC_02503-rpsQ
SAOUHSC_02504
SAOUHSC_02505-rplP
SAOUHSC_02506-rpsC
SAOUHSC_02507-rplV
SAOUHSC_02508-rpsS
SAOUHSC_02509-rplB
SAOUHSC_02510-rplW
SAOUHSC_02511-rplD
SAOUHSC_02512a
SAOUHSC_02512-rplC
SAOUHSC_02515

SAOUHSC_02491-secY
SAOUHSC_02492-rplO
SAOUHSC_02493-rpmD
SAOUHSC_02494-rpsE
SAOUHSC_02495-rplR
SAOUHSC_02496-rplF
SAOUHSC_02498-rpsH
SAOUHSC_02499-rpsN
SAOUHSC_02500-rplE
SAOUHSC_02501-rplX
SAOUHSC_02502-rplN
SAOUHSC_02503-rpsQ
SAOUHSC_02504
SAOUHSC_02505-rplP
SAOUHSC_02506-rpsC
SAOUHSC_02507-rplV
SAOUHSC_02508-rpsS
SAOUHSC_02509-rplB
SAOUHSC_02510-rplW
SAOUHSC_02511-rplD
SAOUHSC_02512a
SAOUHSC_02512-rplC
SAOUHSC_02515
SAOUHSC_02528
SAOUHSC_02612
SAOUHSC_02623
SAOUHSC_02793-pgcA
SAOUHSC_03053-trmE
SAOUHSC_03054-rnpA
SAOUHSC_A01514

SAOUHSC_02491-secY
SAOUHSC_02492-rplO
SAOUHSC_02493-rpmD
SAOUHSC_02494-rpsE
SAOUHSC_02495-rplR
SAOUHSC_02496-rplF
SAOUHSC_02498-rpsH
SAOUHSC_02499-rpsN
SAOUHSC_02500-rplE
SAOUHSC_02501-rplX
SAOUHSC_02502-rplN
SAOUHSC_02503-rpsQ
SAOUHSC_02504
SAOUHSC_02505-rplP
SAOUHSC_02506-rpsC
SAOUHSC_02507-rplV
SAOUHSC_02508-rpsS
SAOUHSC_02509-rplB
SAOUHSC_02510-rplW
SAOUHSC_02511-rplD
SAOUHSC_02512a
SAOUHSC_02512-rplC
SAOUHSC_02515
SAOUHSC_02528
SAOUHSC_02612
SAOUHSC_02623
SAOUHSC_02793-pgcA
SAOUHSC_03052
SAOUHSC_03053-trmE
SAOUHSC_03054-rnpA
SAOUHSC_03055-rpmH
SAOUHSC_A01514

SAOUHSC_02528
SAOUHSC_02612
SAOUHSC_02623
SAOUHSC_02805
SAOUHSC_03053-trmE
SAOUHSC_03054-rnpA

E. Essential genes in three studies

Valentino Essential Genes	Chaudhuri Essential Genes	Santiago Essential Genes
SAOUHSC_00001	SAOUHSC_00001	SAOUHSC_00001
SAOUHSC_00002	SAOUHSC_00002	SAOUHSC_00005
SAOUHSC_00005	SAOUHSC_00003	SAOUHSC_00006
SAOUHSC_00006	SAOUHSC_00005	SAOUHSC_00009
SAOUHSC_00009	SAOUHSC_00006	SAOUHSC_00015
SAOUHSC_00017	SAOUHSC_00015	SAOUHSC_00017
SAOUHSC_00018	SAOUHSC_00018	SAOUHSC_00018
SAOUHSC_00020	SAOUHSC_00020	SAOUHSC_00021
SAOUHSC_00021	SAOUHSC_00021	SAOUHSC_00226
SAOUHSC_00038	SAOUHSC_00226	SAOUHSC_00349
SAOUHSC_00112	SAOUHSC_00336	SAOUHSC_00350
SAOUHSC_00223	SAOUHSC_00345	SAOUHSC_00374
SAOUHSC_00226	SAOUHSC_00349	SAOUHSC_00375
SAOUHSC_00227	SAOUHSC_00350	SAOUHSC_00413
SAOUHSC_00275	SAOUHSC_00375	SAOUHSC_00442
SAOUHSC_00345	SAOUHSC_00442	SAOUHSC_00444
SAOUHSC_00348	SAOUHSC_00444	SAOUHSC_00445
SAOUHSC_00374	SAOUHSC_00454	SAOUHSC_00451
SAOUHSC_00375	SAOUHSC_00461	SAOUHSC_00454
SAOUHSC_00444	SAOUHSC_00474	SAOUHSC_00461
SAOUHSC_00451	SAOUHSC_00475	SAOUHSC_00471
SAOUHSC_00453	SAOUHSC_00482	SAOUHSC_00472
SAOUHSC_00454	SAOUHSC_00484	SAOUHSC_00481
SAOUHSC_00461	SAOUHSC_00489	SAOUHSC_00482
SAOUHSC_00471	SAOUHSC_00490	SAOUHSC_00484
SAOUHSC_00472	SAOUHSC_00491	SAOUHSC_00485
SAOUHSC_00473	SAOUHSC_00493	SAOUHSC_00486
SAOUHSC_00475	SAOUHSC_00509	SAOUHSC_00490
SAOUHSC_00481	SAOUHSC_00510	SAOUHSC_00491
SAOUHSC_00482	SAOUHSC_00511	SAOUHSC_00493
SAOUHSC_00484	SAOUHSC_00516	SAOUHSC_00510
SAOUHSC_00486	SAOUHSC_00518	SAOUHSC_00511
SAOUHSC_00489	SAOUHSC_00519	SAOUHSC_00518
SAOUHSC_00490	SAOUHSC_00520	SAOUHSC_00519
SAOUHSC_00491	SAOUHSC_00521	SAOUHSC_00520
SAOUHSC_00509	SAOUHSC_00525	SAOUHSC_00521
SAOUHSC_00510	SAOUHSC_00527	SAOUHSC_00524
SAOUHSC_00516	SAOUHSC_00529	SAOUHSC_00525

SAOUHSC_00518	SAOUHSC_00530	SAOUHSC_00527
SAOUHSC_00519	SAOUHSC_00549	SAOUHSC_00528
SAOUHSC_00521	SAOUHSC_00575	SAOUHSC_00529
SAOUHSC_00524	SAOUHSC_00577	SAOUHSC_00530
SAOUHSC_00525	SAOUHSC_00578	SAOUHSC_00574
SAOUHSC_00526	SAOUHSC_00579	SAOUHSC_00575
SAOUHSC_00527	SAOUHSC_00611	SAOUHSC_00578
SAOUHSC_00529	SAOUHSC_00640	SAOUHSC_00579
SAOUHSC_00549	SAOUHSC_00642	SAOUHSC_00610
SAOUHSC_00573	SAOUHSC_00643	SAOUHSC_00611
SAOUHSC_00574	SAOUHSC_00645	SAOUHSC_00641
SAOUHSC_00578	SAOUHSC_00741	SAOUHSC_00643
SAOUHSC_00579	SAOUHSC_00742	SAOUHSC_00645
SAOUHSC_00611	SAOUHSC_00752	SAOUHSC_00742
SAOUHSC_00640	SAOUHSC_00760	SAOUHSC_00743
SAOUHSC_00641	SAOUHSC_00762	SAOUHSC_00752
SAOUHSC_00642	SAOUHSC_00769	SAOUHSC_00769
SAOUHSC_00643	SAOUHSC_00771	SAOUHSC_00771
SAOUHSC_00680	SAOUHSC_00781	SAOUHSC_00778
SAOUHSC_00728	SAOUHSC_00785	SAOUHSC_00785
SAOUHSC_00735	SAOUHSC_00788	SAOUHSC_00788
SAOUHSC_00741	SAOUHSC_00790	SAOUHSC_00795
SAOUHSC_00778	SAOUHSC_00803	SAOUHSC_00796
SAOUHSC_00790	SAOUHSC_00849	SAOUHSC_00797
SAOUHSC_00794	SAOUHSC_00850	SAOUHSC_00798
SAOUHSC_00795	SAOUHSC_00851	SAOUHSC_00799
SAOUHSC_00796	SAOUHSC_00868	SAOUHSC_00804
SAOUHSC_00799	SAOUHSC_00870	SAOUHSC_00832
SAOUHSC_00851	SAOUHSC_00920	SAOUHSC_00848
SAOUHSC_00857	SAOUHSC_00921	SAOUHSC_00849
SAOUHSC_00871	SAOUHSC_00947	SAOUHSC_00850
SAOUHSC_00872	SAOUHSC_00954	SAOUHSC_00851
SAOUHSC_00886	SAOUHSC_00998	SAOUHSC_00868
SAOUHSC_00887	SAOUHSC_01028	SAOUHSC_00869
SAOUHSC_00896	SAOUHSC_01038	SAOUHSC_00870
SAOUHSC_00900	SAOUHSC_01040	SAOUHSC_00871
SAOUHSC_00920	SAOUHSC_01050	SAOUHSC_00872
SAOUHSC_00921	SAOUHSC_01063	SAOUHSC_00884
SAOUHSC_00933	SAOUHSC_01075	SAOUHSC_00885
SAOUHSC_00934	SAOUHSC_01077	SAOUHSC_00886
SAOUHSC_00938	SAOUHSC_01078	SAOUHSC_00887
SAOUHSC_00943	SAOUHSC_01093	SAOUHSC_00888

SAOUHSC_01023	SAOUHSC_01148	SAOUHSC_00889
SAOUHSC_01028	SAOUHSC_01149	SAOUHSC_00900
SAOUHSC_01035	SAOUHSC_01150	SAOUHSC_00903
SAOUHSC_01063	SAOUHSC_01159	SAOUHSC_00904
SAOUHSC_01075	SAOUHSC_01176	SAOUHSC_00905
SAOUHSC_01093	SAOUHSC_01179	SAOUHSC_00921
SAOUHSC_01106	SAOUHSC_01183	SAOUHSC_00933
SAOUHSC_01118	SAOUHSC_01188	SAOUHSC_00943
SAOUHSC_01143	SAOUHSC_01190	SAOUHSC_00947
SAOUHSC_01144	SAOUHSC_01191	SAOUHSC_00953
SAOUHSC_01145	SAOUHSC_01197	SAOUHSC_00980
SAOUHSC_01146	SAOUHSC_01198	SAOUHSC_00981
SAOUHSC_01154	SAOUHSC_01208	SAOUHSC_00982
SAOUHSC_01176	SAOUHSC_01211	SAOUHSC_00983
SAOUHSC_01178	SAOUHSC_01214	SAOUHSC_00985
SAOUHSC_01179	SAOUHSC_01216	SAOUHSC_01000
SAOUHSC_01183	SAOUHSC_01222	SAOUHSC_01002
SAOUHSC_01190	SAOUHSC_01234	SAOUHSC_01035
SAOUHSC_01191	SAOUHSC_01235	SAOUHSC_01036
SAOUHSC_01196	SAOUHSC_01236	SAOUHSC_01040
SAOUHSC_01197	SAOUHSC_01237	SAOUHSC_01041
SAOUHSC_01198	SAOUHSC_01238	SAOUHSC_01042
SAOUHSC_01199	SAOUHSC_01240	SAOUHSC_01043
SAOUHSC_01200	SAOUHSC_01241	SAOUHSC_01063
SAOUHSC_01203	SAOUHSC_01244	SAOUHSC_01075
SAOUHSC_01207	SAOUHSC_01246	SAOUHSC_01077
SAOUHSC_01224	SAOUHSC_01260	SAOUHSC_01078
SAOUHSC_01232	SAOUHSC_01287	SAOUHSC_01092
SAOUHSC_01234	SAOUHSC_01337	SAOUHSC_01093
SAOUHSC_01235	SAOUHSC_01350	SAOUHSC_01142
SAOUHSC_01236	SAOUHSC_01351	SAOUHSC_01143
SAOUHSC_01237	SAOUHSC_01352	SAOUHSC_01144
SAOUHSC_01238	SAOUHSC_01359	SAOUHSC_01145
SAOUHSC_01239	SAOUHSC_01361	SAOUHSC_01146
SAOUHSC_01240	SAOUHSC_01362	SAOUHSC_01147
SAOUHSC_01241	SAOUHSC_01373	SAOUHSC_01148
SAOUHSC_01259	SAOUHSC_01467	SAOUHSC_01149
SAOUHSC_01260	SAOUHSC_01470	SAOUHSC_01150
SAOUHSC_01262	SAOUHSC_01473	SAOUHSC_01158
SAOUHSC_01270	SAOUHSC_01477	SAOUHSC_01159
SAOUHSC_01285	SAOUHSC_01490	SAOUHSC_01163
SAOUHSC_01287	SAOUHSC_01492	SAOUHSC_01166

SAOUHSC_01338	SAOUHSC_01598	SAOUHSC_01168
SAOUHSC_01350	SAOUHSC_01599	SAOUHSC_01169
SAOUHSC_01374	SAOUHSC_01625	SAOUHSC_01170
SAOUHSC_01390	SAOUHSC_01627	SAOUHSC_01171
SAOUHSC_01409	SAOUHSC_01661	SAOUHSC_01172
SAOUHSC_01434	SAOUHSC_01666	SAOUHSC_01178
SAOUHSC_01435	SAOUHSC_01668	SAOUHSC_01179
SAOUHSC_01441	SAOUHSC_01672	SAOUHSC_01183
SAOUHSC_01473	SAOUHSC_01683	SAOUHSC_01188
SAOUHSC_01474	SAOUHSC_01684	SAOUHSC_01189
SAOUHSC_01487	SAOUHSC_01700	SAOUHSC_01190
SAOUHSC_01490	SAOUHSC_01701	SAOUHSC_01194
SAOUHSC_01492	SAOUHSC_01714	SAOUHSC_01197
SAOUHSC_01496	SAOUHSC_01720	SAOUHSC_01198
SAOUHSC_01501	SAOUHSC_01721	SAOUHSC_01199
SAOUHSC_01550	SAOUHSC_01725	SAOUHSC_01200
SAOUHSC_01551	SAOUHSC_01726	SAOUHSC_01201
SAOUHSC_01554	SAOUHSC_01727	SAOUHSC_01205
SAOUHSC_01555	SAOUHSC_01737	SAOUHSC_01207
SAOUHSC_01573	SAOUHSC_01738	SAOUHSC_01209
SAOUHSC_01619	SAOUHSC_01755	SAOUHSC_01210
SAOUHSC_01623	SAOUHSC_01756	SAOUHSC_01211
SAOUHSC_01624	SAOUHSC_01757	SAOUHSC_01216
SAOUHSC_01625	SAOUHSC_01766	SAOUHSC_01218
SAOUHSC_01662	SAOUHSC_01770	SAOUHSC_01222
SAOUHSC_01663	SAOUHSC_01777	SAOUHSC_01228
SAOUHSC_01669	SAOUHSC_01784	SAOUHSC_01230
SAOUHSC_01672	SAOUHSC_01785	SAOUHSC_01234
SAOUHSC_01682	SAOUHSC_01786	SAOUHSC_01235
SAOUHSC_01683	SAOUHSC_01787	SAOUHSC_01236
SAOUHSC_01684	SAOUHSC_01788	SAOUHSC_01238
SAOUHSC_01687	SAOUHSC_01791	SAOUHSC_01239
SAOUHSC_01698	SAOUHSC_01806	SAOUHSC_01240
SAOUHSC_01700	SAOUHSC_01807	SAOUHSC_01241
SAOUHSC_01727	SAOUHSC_01837	SAOUHSC_01243
SAOUHSC_01737	SAOUHSC_01839	SAOUHSC_01244
SAOUHSC_01738	SAOUHSC_01856	SAOUHSC_01245
SAOUHSC_01741	SAOUHSC_01866	SAOUHSC_01246
SAOUHSC_01746	SAOUHSC_01871	SAOUHSC_01249
SAOUHSC_01753	SAOUHSC_01875	SAOUHSC_01250
SAOUHSC_01755	SAOUHSC_01908	SAOUHSC_01252
SAOUHSC_01756	SAOUHSC_01909	SAOUHSC_01259

SAOUHSC_01757	SAOUHSC_01928	SAOUHSC_01260
SAOUHSC_01760	SAOUHSC_01930	SAOUHSC_01262
SAOUHSC_01767	SAOUHSC_02102	SAOUHSC_01265
SAOUHSC_01772	SAOUHSC_02106	SAOUHSC_01267
SAOUHSC_01773	SAOUHSC_02107	SAOUHSC_01287
SAOUHSC_01774	SAOUHSC_02114	SAOUHSC_01337
SAOUHSC_01777	SAOUHSC_02117	SAOUHSC_01351
SAOUHSC_01778	SAOUHSC_02118	SAOUHSC_01352
SAOUHSC_01784	SAOUHSC_02122	SAOUHSC_01373
SAOUHSC_01785	SAOUHSC_02123	SAOUHSC_01374
SAOUHSC_01786	SAOUHSC_02132	SAOUHSC_01424
SAOUHSC_01787	SAOUHSC_02133	SAOUHSC_01434
SAOUHSC_01788	SAOUHSC_02140	SAOUHSC_01467
SAOUHSC_01789	SAOUHSC_02151	SAOUHSC_01470
SAOUHSC_01791	SAOUHSC_02152	SAOUHSC_01473
SAOUHSC_01792	SAOUHSC_02254	SAOUHSC_01474
SAOUHSC_01804	SAOUHSC_02260	SAOUHSC_01481
SAOUHSC_01805	SAOUHSC_02277	SAOUHSC_01482
SAOUHSC_01807	SAOUHSC_02279	SAOUHSC_01483
SAOUHSC_01808	SAOUHSC_02280	SAOUHSC_01487
SAOUHSC_01809	SAOUHSC_02306	SAOUHSC_01488
SAOUHSC_01811	SAOUHSC_02317	SAOUHSC_01491
SAOUHSC_01827	SAOUHSC_02318	SAOUHSC_01492
SAOUHSC_01829	SAOUHSC_02327	SAOUHSC_01496
SAOUHSC_01837	SAOUHSC_02336	SAOUHSC_01501
SAOUHSC_01839	SAOUHSC_02337	SAOUHSC_01591
SAOUHSC_01841	SAOUHSC_02357	SAOUHSC_01592
SAOUHSC_01856	SAOUHSC_02359	SAOUHSC_01598
SAOUHSC_01857	SAOUHSC_02361	SAOUHSC_01599
SAOUHSC_01858	SAOUHSC_02366	SAOUHSC_01612
SAOUHSC_01859	SAOUHSC_02368	SAOUHSC_01613
SAOUHSC_01871	SAOUHSC_02371	SAOUHSC_01614
SAOUHSC_01875	SAOUHSC_02399	SAOUHSC_01619
SAOUHSC_01905	SAOUHSC_02405	SAOUHSC_01623
SAOUHSC_01906	SAOUHSC_02407	SAOUHSC_01624
SAOUHSC_01908	SAOUHSC_02477	SAOUHSC_01635
SAOUHSC_01909	SAOUHSC_02478	SAOUHSC_01659
SAOUHSC_01911	SAOUHSC_02484	SAOUHSC_01661
SAOUHSC_01924	SAOUHSC_02485	SAOUHSC_01662
SAOUHSC_01961	SAOUHSC_02486	SAOUHSC_01663
SAOUHSC_01962	SAOUHSC_02487	SAOUHSC_01666
SAOUHSC_01973	SAOUHSC_02488	SAOUHSC_01668

SAOUHSC_01977	SAOUHSC_02489	SAOUHSC_01670
SAOUHSC_01993	SAOUHSC_02490	SAOUHSC_01672
SAOUHSC_02084	SAOUHSC_02499	SAOUHSC_01690
SAOUHSC_02102	SAOUHSC_02500	SAOUHSC_01697
SAOUHSC_00015	SAOUHSC_00009	SAOUHSC_01698
SAOUHSC_00022	SAOUHSC_00223	SAOUHSC_01699
SAOUHSC_00023	SAOUHSC_00225	SAOUHSC_01700
SAOUHSC_00100	SAOUHSC_00227	SAOUHSC_01701
SAOUHSC_00225	SAOUHSC_00348	SAOUHSC_01720
SAOUHSC_00349	SAOUHSC_00451	SAOUHSC_01721
SAOUHSC_00396	SAOUHSC_00471	SAOUHSC_01722
SAOUHSC_00442	SAOUHSC_00472	SAOUHSC_01727
SAOUHSC_00493	SAOUHSC_00524	SAOUHSC_01738
SAOUHSC_00511	SAOUHSC_00528	SAOUHSC_01741
SAOUHSC_00520	SAOUHSC_00574	SAOUHSC_01746
SAOUHSC_00528	SAOUHSC_00620	SAOUHSC_01750
SAOUHSC_00530	SAOUHSC_00641	SAOUHSC_01753
SAOUHSC_00577	SAOUHSC_00728	SAOUHSC_01756
SAOUHSC_00610	SAOUHSC_00743	SAOUHSC_01757
SAOUHSC_00742	SAOUHSC_00793	SAOUHSC_01766
SAOUHSC_00743	SAOUHSC_00795	SAOUHSC_01767
SAOUHSC_00752	SAOUHSC_00796	SAOUHSC_01772
SAOUHSC_00762	SAOUHSC_00797	SAOUHSC_01773
SAOUHSC_00769	SAOUHSC_00798	SAOUHSC_01774
SAOUHSC_00771	SAOUHSC_00799	SAOUHSC_01776
SAOUHSC_00781	SAOUHSC_00804	SAOUHSC_01778
SAOUHSC_00785	SAOUHSC_00847	SAOUHSC_01782
SAOUHSC_00786	SAOUHSC_00848	SAOUHSC_01783
SAOUHSC_00797	SAOUHSC_00869	SAOUHSC_01784
SAOUHSC_00804	SAOUHSC_00871	SAOUHSC_01785
SAOUHSC_00819	SAOUHSC_00872	SAOUHSC_01786
SAOUHSC_00847	SAOUHSC_00881	SAOUHSC_01787
SAOUHSC_00848	SAOUHSC_00892	SAOUHSC_01789
SAOUHSC_00849	SAOUHSC_00900	SAOUHSC_01791
SAOUHSC_00850	SAOUHSC_00903	SAOUHSC_01792
SAOUHSC_00867	SAOUHSC_00922	SAOUHSC_01793
SAOUHSC_00868	SAOUHSC_00933	SAOUHSC_01807
SAOUHSC_00869	SAOUHSC_00934	SAOUHSC_01809
SAOUHSC_00870	SAOUHSC_00943	SAOUHSC_01811
SAOUHSC_00884	SAOUHSC_00957	SAOUHSC_01829
SAOUHSC_00885	SAOUHSC_00980	SAOUHSC_01839
SAOUHSC_00888	SAOUHSC_01035	SAOUHSC_01856

SAOUHSC_00889	SAOUHSC_01036	SAOUHSC_01859
SAOUHSC_00939	SAOUHSC_01092	SAOUHSC_01866
SAOUHSC_00947	SAOUHSC_01100	SAOUHSC_01871
SAOUHSC_00954	SAOUHSC_01106	SAOUHSC_01875
SAOUHSC_00957	SAOUHSC_01119	SAOUHSC_01915
SAOUHSC_00980	SAOUHSC_01144	SAOUHSC_01916
SAOUHSC_00981	SAOUHSC_01145	SAOUHSC_01961
SAOUHSC_00985	SAOUHSC_01146	SAOUHSC_01962
SAOUHSC_01003	SAOUHSC_01147	SAOUHSC_01973
SAOUHSC_01036	SAOUHSC_01154	SAOUHSC_02106
SAOUHSC_01092	SAOUHSC_01178	SAOUHSC_02107
SAOUHSC_01142	SAOUHSC_01189	SAOUHSC_02114
SAOUHSC_01147	SAOUHSC_01199	SAOUHSC_02116
SAOUHSC_01148	SAOUHSC_01201	SAOUHSC_02117
SAOUHSC_01149	SAOUHSC_01205	SAOUHSC_02118
SAOUHSC_01150	SAOUHSC_01207	SAOUHSC_02122
SAOUHSC_01158	SAOUHSC_01209	SAOUHSC_02123
SAOUHSC_01159	SAOUHSC_01210	SAOUHSC_02133
SAOUHSC_01189	SAOUHSC_01232	SAOUHSC_02140
SAOUHSC_01201	SAOUHSC_01243	SAOUHSC_02143
SAOUHSC_01205	SAOUHSC_01245	SAOUHSC_02151
SAOUHSC_01209	SAOUHSC_01249	SAOUHSC_02152
SAOUHSC_01214	SAOUHSC_01250	SAOUHSC_02255
SAOUHSC_01222	SAOUHSC_01252	SAOUHSC_02277
SAOUHSC_01230	SAOUHSC_01285	SAOUHSC_02278
SAOUHSC_01233	SAOUHSC_01333	SAOUHSC_02279
SAOUHSC_01243	SAOUHSC_01374	SAOUHSC_02280
SAOUHSC_01244	SAOUHSC_01424	SAOUHSC_02318
SAOUHSC_01245	SAOUHSC_01434	SAOUHSC_02336
SAOUHSC_01246	SAOUHSC_01435	SAOUHSC_02357
SAOUHSC_01249	SAOUHSC_01462	SAOUHSC_02359
SAOUHSC_01252	SAOUHSC_01466	SAOUHSC_02368
SAOUHSC_01263	SAOUHSC_01474	SAOUHSC_02399
SAOUHSC_01289	SAOUHSC_01496	SAOUHSC_02407
SAOUHSC_01293	SAOUHSC_01501	SAOUHSC_02478
SAOUHSC_01294	SAOUHSC_01504	SAOUHSC_02482
SAOUHSC_01333	SAOUHSC_01592	SAOUHSC_02485
SAOUHSC_01351	SAOUHSC_01605	SAOUHSC_02486
SAOUHSC_01362	SAOUHSC_01623	SAOUHSC_02487
SAOUHSC_01373	SAOUHSC_01624	SAOUHSC_02488
SAOUHSC_01410	SAOUHSC_01662	SAOUHSC_02489
SAOUHSC_01424	SAOUHSC_01663	SAOUHSC_02490

SAOUHSC_01467 SAOUHSC_01678 SAOUHSC_02491
SAOUHSC_01470 SAOUHSC_01682 SAOUHSC_02492
SAOUHSC_01477 SAOUHSC_01690 SAOUHSC_02493
SAOUHSC_01483 SAOUHSC_01697 SAOUHSC_02494
SAOUHSC_01544 SAOUHSC_01722 SAOUHSC_02495
SAOUHSC_01578 SAOUHSC_01739 SAOUHSC_02496
SAOUHSC_01579 SAOUHSC_01741 SAOUHSC_02498
SAOUHSC_01580 SAOUHSC_01742 SAOUHSC_02499
SAOUHSC_01592 SAOUHSC_01746 SAOUHSC_02500
SAOUHSC_01598 SAOUHSC_01750 SAOUHSC_02501
SAOUHSC_01599 SAOUHSC_01751 SAOUHSC_02502
SAOUHSC_01605 SAOUHSC_01753 SAOUHSC_02503
SAOUHSC_01627 SAOUHSC_01767 SAOUHSC_02504
SAOUHSC_01666 SAOUHSC_01782 SAOUHSC_02505
SAOUHSC_01690 SAOUHSC_01792 SAOUHSC_02506
SAOUHSC_01697 SAOUHSC_01795 SAOUHSC_02507
SAOUHSC_01701 SAOUHSC_01808 SAOUHSC_02508
SAOUHSC_01714 SAOUHSC_01809 SAOUHSC_02509
SAOUHSC_01722 SAOUHSC_01811 SAOUHSC_02510
SAOUHSC_01725 SAOUHSC_01827 SAOUHSC_02511
SAOUHSC_01726 SAOUHSC_01829 SAOUHSC_02512
SAOUHSC_01766 SAOUHSC_01979 SAOUHSC_02512a
SAOUHSC_01776 SAOUHSC_02116 SAOUHSC_02527
SAOUHSC_01795 SAOUHSC_02255 SAOUHSC_02612
SAOUHSC_02002 SAOUHSC_02491 SAOUHSC_02623
SAOUHSC_02017 SAOUHSC_02492 SAOUHSC_02793
SAOUHSC_02053 SAOUHSC_02493 SAOUHSC_03052
SAOUHSC_02054 SAOUHSC_02494 SAOUHSC_03053
SAOUHSC_02055 SAOUHSC_02495 SAOUHSC_03054
SAOUHSC_02065 SAOUHSC_02496 SAOUHSC_03055
SAOUHSC_02076 SAOUHSC_02498 SAOUHSC_A01514
SAOUHSC_02106 SAOUHSC_02501
SAOUHSC_02107 SAOUHSC_02502
SAOUHSC_02114 SAOUHSC_02503
SAOUHSC_02116 SAOUHSC_02504
SAOUHSC_02117 SAOUHSC_02505
SAOUHSC_02118 SAOUHSC_02506
SAOUHSC_02122 SAOUHSC_02507
SAOUHSC_02123 SAOUHSC_02508
SAOUHSC_02132 SAOUHSC_02509
SAOUHSC_02133 SAOUHSC_02510
SAOUHSC_02140 SAOUHSC_02511

SAOUHSC_02151   SAOUHSC_02512
SAOUHSC_02152   SAOUHSC_02527
SAOUHSC_02154   SAOUHSC_02571
SAOUHSC_02164   SAOUHSC_02572
SAOUHSC_02175   SAOUHSC_02575
SAOUHSC_02183   SAOUHSC_02612
SAOUHSC_02209   SAOUHSC_02623
SAOUHSC_02214   SAOUHSC_02720
SAOUHSC_02215   SAOUHSC_02757
SAOUHSC_02219   SAOUHSC_02791
SAOUHSC_02224   SAOUHSC_02805
SAOUHSC_02235   SAOUHSC_02859
SAOUHSC_02237   SAOUHSC_02860
SAOUHSC_02238   SAOUHSC_03049
SAOUHSC_02254   SAOUHSC_03052
SAOUHSC_02255   SAOUHSC_03053
SAOUHSC_02277   SAOUHSC_03054
SAOUHSC_02279   SAOUHSC_03055
SAOUHSC_02280
SAOUHSC_02294
SAOUHSC_02306
SAOUHSC_02307
SAOUHSC_02317
SAOUHSC_02318
SAOUHSC_02325
SAOUHSC_02327
SAOUHSC_02332
SAOUHSC_02336
SAOUHSC_02357
SAOUHSC_02359
SAOUHSC_02361
SAOUHSC_02368
SAOUHSC_02371
SAOUHSC_02379
SAOUHSC_02383
SAOUHSC_02399
SAOUHSC_02405
SAOUHSC_02407
SAOUHSC_02410
SAOUHSC_02411
SAOUHSC_02412
SAOUHSC_02416

SAOUHSC_02437
SAOUHSC_02438
SAOUHSC_02440
SAOUHSC_02477
SAOUHSC_02478
SAOUHSC_02484
SAOUHSC_02485
SAOUHSC_02486
SAOUHSC_02487
SAOUHSC_02488
SAOUHSC_02489
SAOUHSC_02490
SAOUHSC_02491
SAOUHSC_02492
SAOUHSC_02493
SAOUHSC_02494
SAOUHSC_02495
SAOUHSC_02496
SAOUHSC_02498
SAOUHSC_02499
SAOUHSC_02500
SAOUHSC_02501
SAOUHSC_02502
SAOUHSC_02503
SAOUHSC_02504
SAOUHSC_02505
SAOUHSC_02506
SAOUHSC_02507
SAOUHSC_02508
SAOUHSC_02509
SAOUHSC_02510
SAOUHSC_02511
SAOUHSC_02512
SAOUHSC_02527
SAOUHSC_02534
SAOUHSC_02571
SAOUHSC_02623
SAOUHSC_02707
SAOUHSC_02720
SAOUHSC_02721
SAOUHSC_02745
SAOUHSC_02860

SAOUHSC_02872
SAOUHSC_03043
SAOUHSC_03054
SAOUHSC_03055

F. Genes important for *S. aureus* growth at different temperatures

16°C blunt

gene locus	corrected p-value	ratio experiment/control
SAOUHSC_01857	6.88E-05	6.06

23°C blunt

gene locus	corrected p-value	ratio experiment/control
SAOUHSC_01857	1.69E-04	5.03

37°C blunt

gene locus	corrected p-value	ratio experiment/control
SAOUHSC_00536-ilvE	6.18E-05	14.59
SAOUHSC_01618-ispA	9.63E-04	12.82
SAOUHSC_01613-bkdA1	2.90E-03	8.00
SAOUHSC_01611	2.03E-04	6.50
SAOUHSC_00525	5.59E-03	6.40
SAOUHSC_01612-bkdA2	2.36E-02	6.18
SAOUHSC_01153	5.05E-03	0.08
SAOUHSC_01154-sepF	3.55E-02	0.07

43°C blunt

gene locus	corrected p-value	ratio experiment/control
SAOUHSC_00014	2.41E-03	22.85
SAOUHSC_00015	1.60E-03	0.13
SAOUHSC_00022-yycH	2.20E-04	0.03
SAOUHSC_00023-yycI	7.71E-05	0.03
SAOUHSC_00290	1.27E-08	0.02
SAOUHSC_00364	7.20E-06	0.07
SAOUHSC_00373	4.35E-05	10.01
SAOUHSC_00420	1.51E-05	10.69
SAOUHSC_00455	1.81E-02	0.17
SAOUHSC_00462	4.95E-03	6.42
SAOUHSC_00464-ksgA	8.11E-04	8.97
SAOUHSC_00488	2.15E-04	0.04
SAOUHSC_00504-yakI	4.53E-07	0.02

SAOUHSC_00505-clpC	5.43E-03	0.14
SAOUHSC_00618	3.44E-03	0.07
SAOUHSC_00646-pbp4	2.64E-02	0.15
SAOUHSC_00668-vraG	2.89E-04	0.19
SAOUHSC_00675	2.20E-02	0.16
SAOUHSC_00678	1.29E-08	0.02
SAOUHSC_00718-718	4.58E-03	0.06
SAOUHSC_00755	1.67E-02	22.75
SAOUHSC_00756	1.18E-09	0.02
SAOUHSC_00760-gdpS	1.66E-07	0.04
SAOUHSC_00794	1.14E-03	0.02
SAOUHSC_00803	7.71E-05	0.07
SAOUHSC_00892	8.05E-03	0.11
SAOUHSC_00935	1.02E-02	5.90
SAOUHSC_00965-CAAX	3.82E-02	5.04
SAOUHSC_00982-menF	1.59E-05	13.25
SAOUHSC_00983-menD	1.03E-09	9.73
SAOUHSC_00996	1.92E-02	0.06
SAOUHSC_00997-lcpB	2.60E-03	0.17
SAOUHSC_01025-1025	2.01E-09	0.04
SAOUHSC_01039	4.45E-04	20.47
SAOUHSC_01050-1050	4.37E-09	0.03
SAOUHSC_01095	4.91E-02	0.07
SAOUHSC_01153	2.50E-05	0.04
SAOUHSC_01154-sepF	1.79E-02	0.10
SAOUHSC_01165	2.23E-05	0.06
SAOUHSC_01184	7.77E-04	0.12
SAOUHSC_01186	1.60E-03	0.04
SAOUHSC_01203-rnc	3.82E-02	0.16
SAOUHSC_01265	9.22E-04	8.69
SAOUHSC_01269	1.50E-07	0.03
SAOUHSC_01271	1.93E-05	0.02
SAOUHSC_01278	1.12E-10	15.60
SAOUHSC_01346	6.44E-04	5.45
SAOUHSC_01359-mprF	1.10E-21	0.05
SAOUHSC_01361-lcpA	6.76E-03	0.04
SAOUHSC_01427	7.39E-10	25.24

SAOUHSC_01437	6.33E-04	0.10
SAOUHSC_01462-gpsB	2.03E-02	0.03
SAOUHSC_01480	3.26E-04	10.79
SAOUHSC_01481	3.71E-02	5.98
SAOUHSC_01482-aroB	2.57E-08	20.95
SAOUHSC_01483-aroF	2.36E-03	8.05
SAOUHSC_01497	2.16E-04	0.07
SAOUHSC_01585	1.48E-06	8.10
SAOUHSC_01586	7.43E-08	32.16
SAOUHSC_01618-ispA	2.41E-04	20.18
SAOUHSC_01627	6.59E-04	0.04
SAOUHSC_01650	3.17E-03	0.06
SAOUHSC_01652-pbp3	1.10E-11	0.02
SAOUHSC_01656	2.90E-07	12.64
SAOUHSC_01657	7.87E-04	12.92
SAOUHSC_01682-dnaJ	2.40E-03	0.04
SAOUHSC_01683-dnaK	1.90E-09	0.02
SAOUHSC_01685-hrcA	1.88E-08	0.02
SAOUHSC_01696	1.11E-04	0.19
SAOUHSC_01699-aroE	9.58E-03	8.35
SAOUHSC_01739	3.49E-02	0.09
SAOUHSC_01758-mreD	1.27E-05	0.02
SAOUHSC_01759-mreC	5.64E-07	0.02
SAOUHSC_01779-tig	3.24E-05	16.96
SAOUHSC_01803	7.76E-16	0.02
SAOUHSC_01810	2.94E-04	0.06
SAOUHSC_01821	3.24E-05	17.85
SAOUHSC_01827-ezrA	1.15E-03	0.07
SAOUHSC_01852-aroA	3.55E-04	11.38
SAOUHSC_01854	1.59E-05	0.04
SAOUHSC_01857	2.98E-09	0.16
SAOUHSC_01858	2.20E-04	0.02
SAOUHSC_01895-lytD	3.18E-03	10.48
SAOUHSC_01913	2.12E-02	0.04
SAOUHSC_01960-hemY	2.04E-05	9.79
SAOUHSC_01975-sbcD	4.58E-03	6.24

SAOUHSC_01979	4.06E-04	35.37
SAOUHSC_01997	4.52E-02	0.10
SAOUHSC_02004	7.92E-05	0.05
SAOUHSC_02012-sgtB	4.58E-03	6.77
SAOUHSC_02121	4.80E-05	0.06
SAOUHSC_02131	3.71E-02	0.13
SAOUHSC_02143	7.16E-05	19.13
SAOUHSC_02264	9.53E-04	0.20
SAOUHSC_02268	6.06E-03	0.15
SAOUHSC_02305-alr1	6.74E-07	0.08
SAOUHSC_02308	2.33E-07	0.03
SAOUHSC_02309	9.60E-04	0.13
SAOUHSC_02319-rodA	1.32E-09	0.02
SAOUHSC_02337-murA	1.48E-06	0.02
SAOUHSC_02341-F-type	2.39E-04	0.07
SAOUHSC_02345-F-type	1.33E-05	0.05
SAOUHSC_02347-F-type	1.58E-02	0.17
SAOUHSC_02350-F-type	1.66E-02	0.06
SAOUHSC_02351	1.79E-02	0.05
SAOUHSC_02358	4.01E-03	23.16
SAOUHSC_02366	3.31E-04	23.54
SAOUHSC_02369	3.70E-02	11.09
SAOUHSC_02372	1.26E-02	0.08
SAOUHSC_02383	7.88E-13	0.02
SAOUHSC_02481-cbiQ	5.09E-05	0.01
SAOUHSC_02483-cbiO	1.59E-02	0.11
SAOUHSC_02552	3.66E-05	0.06
SAOUHSC_02571	8.02E-06	0.03
SAOUHSC_02589	9.27E-03	5.47
SAOUHSC_02611-lyrA	3.69E-05	24.09
SAOUHSC_02664	3.30E-08	7.27
SAOUHSC_02690	8.02E-03	11.85
SAOUHSC_02885	5.76E-06	5.96

G. Tn-Seq hits with a panel of β-lactam antibiotics in MW2 and USA300

USA300 0.1 ug/ml oxacillin	ratio	USA300 1 ug/ml oxacillin	ratio	USA300 10 ug/ml oxacillin	ratio
USA300HOU_0413	38.591	USA300HOU_0414	4364.00	USA300HOU_0413	7897.59
USA300HOU_0504	37.224	USA300HOU_2045	2342.00	USA300HOU_0504	7415.57
USA300HOU_1784	33.296	USA300HOU_0413	2129.04	USA300HOU_0414	3492.00
USA300HOU_1197	32.898	USA300HOU_0504	2024.22	USA300HOU_0383	2303.80
USA300HOU_0959	29.970	USA300HOU_0383	581.00	USA300HOU_1670	979.56
USA300HOU_0505	26.869	USA300HOU_1666	170.09	USA300HOU_2045	674.50
USA300HOU_2116	21.058	USA300HOU_1194	70.67	USA300HOU_1632	552.33
USA300HOU_1823	17.964	USA300HOU_2515	34.03	USA300HOU_0275	40.56
USA300HOU_0993	13.532	USA300HOU_0411	24.59	USA300HOU_1052	17.60
USA300HOU_1727	12.620	USA300HOU_0275	12.78		
USA300HOU_1403	12.402	USA300HOU_1667	12.25		
USA300HOU_0990	12.370				
USA300HOU_1539	11.124				
USA300HOU_1404	10.499				
USA300HOU_0136	0.099				
USA300HOU_1385	0.099				
USA300HOU_0035	0.097				
USA300HOU_2368	0.092				
USA300HOU_1293	0.091				
USA300HOU_0693	0.090				
USA300HOU_0703	0.089				
USA300HOU_1856	0.089				
USA300HOU_1765	0.088				
USA300HOU_2655	0.087				
USA300HOU_2130	0.085				
USA300HOU_0718	0.085				
USA300HOU_2503	0.085				
USA300HOU_1729	0.067				
USA300HOU_1728	0.064				
USA300HOU_0498	0.063				
USA300HOU_0351	0.062				
USA300HOU_2302	0.058				
USA300HOU_0681	0.056				
USA300HOU_0769	0.056				
USA300HOU_0662	0.053				
USA300HOU_2109	0.053				
USA300HOU_1350	0.052				
USA300HOU_0682	0.043				
USA300HOU_0683	0.024				
USA300HOU_0794	0.023				
USA300HOU_0031	0.023				

MW2 WT 0.1µg/ml oxacillin	ratio	MW2 WT 1µg/ml oxacillin	ratio	MW2 WT 10µg/ml oxacillin	ratio
aroA	41.88	clpP	9297.33	relA	18665.86
MW0926	14.68	MW2393	3058.90	MW0465	12952.65
		MW0465	3052.45	guaB	401.18
		relA	2661.21	clpP	156.33
		hemB	1359.00	hemB	118.00
		hemE	334.50	MW0177	69.80
		topB	173.81	pgk	33.27
		guaB	140.66		
		aroA	134.63		
		guaA	121.43		
		hemH	70.87		
		menE	59.88		
		aroB	45.41		
		aroC	36.04		
		pyrC	34.52		
		MW0765	16.10		
		hemY	14.92		
		MW1872	13.76		
		MW0913	10.03		

MW2 ∆hsdR 0.1 ug/ml oxacillin	ratio	MW2 ∆hsdR 1 ug/ml oxacillin	ratio		MW2 10 ∆hsdR ug/ml oxacillin	ratio
MW2301	19.64	clpP		7491.25	relA	2188.66
MW0883	14.12	srrA		1121.30	clpP	1029.00
MW0926	14.11	fmt		1112.95	MW0465	943.76
pbuX	10.06	relA		1028.63	guaB	115.70
aapA	0.03	MW0465		425.69	MW2041	67.07
		guaB		194.70		
		MW0883		54.23		
		MW1145		32.00		
		comEB		12.39		
		MW0153		6.18		

USA300 + 8µg/ml mecillinam		USA300 + 0.2µg/ml cefoxitin	
Locus	ratio	Locus	ratio
USA300HOU_0504	48.66	USA300HOU_1869	10.98
USA300HOU_0413	48.04	USA300HOU_1841	10.95
USA300HOU_2715	9.07	USA300HOU_1680	10.51
USA300HOU_2714	8.81	USA300HOU_1403	10.35
USA300HOU_1632	7.00	USA300HOU_1753	10.34
USA300HOU_1667	5.50	USA300HOU_0014	10.31
USA300HOU_1164	4.49	USA300HOU_1784	10.27
USA300HOU_0505	4.26	USA300HOU_0836	10.09
USA300HOU_1699	4.23	USA300HOU_1758	0.10
USA300HOU_0021	4.17	USA300HOU_0995	0.10
USA300HOU_0497	0.19	USA300HOU_1082	0.10
USA300HOU_2302	0.18	USA300HOU_2694	0.10
USA300HOU_1828	0.16	USA300HOU_2332	0.10
USA300HOU_0693	0.15	USA300HOU_2126	0.10
USA300HOU_0662	0.14	USA300HOU_1882	0.09
USA300HOU_1559	0.13	USA300HOU_2702	0.09
USA300HOU_0498	0.12	USA300HOU_1043	0.09
USA300HOU_1022	0.11	USA300HOU_1497	0.09
USA300HOU_0769	0.08	USA300HOU_1873	0.09
USA300HOU_1224	0.07	USA300HOU_0790	0.09
USA300HOU_0683	0.07	USA300HOU_1856	0.09
USA300HOU_0794	0.07	USA300HOU_2598	0.09
USA300HOU_1350	0.07	USA300HOU_1203	0.09
USA300HOU_0681	0.06	USA300HOU_2276	0.09
USA300HOU_0682	0.04	USA300HOU_1933	0.09
USA300HOU_0031	0.03	USA300HOU_0809	0.08
USA300HOU_0482	0.03	USA300HOU_2109	0.08
		USA300HOU_1357	0.08
		USA300HOU_0681	0.08
		USA300HOU_2120	0.08
		USA300HOU_0645	0.08
		USA300HOU_0722	0.08
		USA300HOU_0784	0.08
		USA300HOU_0636	0.07
		USA300HOU_1067	0.07
		USA300HOU_0072	0.07
		USA300HOU_2254	0.07
		USA300HOU_1880	0.07
		USA300HOU_1308	0.06
		USA300HOU_0100	0.06
		USA300HOU_1093	0.06
		USA300HOU_0684	0.06
		USA300HOU_0588	0.06
		USA300HOU_1846	0.05
		USA300HOU_0683	0.05
		USA300HOU_2587	0.05
		USA300HOU_1872	0.05

USA300HOU_1495	0.05
USA300HOU_0693	0.05
USA300HOU_1022	0.05
USA300HOU_1876	0.04
USA300HOU_1859	0.04
USA300HOU_2038	0.04
USA300HOU_0794	0.04
USA300HOU_1913	0.04
USA300HOU_1154	0.04
USA300HOU_2175	0.04
USA300HOU_1224	0.03
USA300HOU_1042	0.03
USA300HOU_2361	0.03
USA300HOU_0480	0.02
USA300HOU_0682	0.02
USA300HOU_1556	0.02
USA300HOU_0485	0.02
USA300HOU_1155	0.02
USA300HOU_2091	0.02
USA300HOU_0889	0.02
USA300HOU_0785	0.01
USA300HOU_0031	0.01
USA300HOU_0482	0.01

MW2 32ug/ml mecillinam		MW2 0.4ug/ml cefoxitin	
Locus	ratio	Locus	ratio
MW1169	164.125	rpsC	130.583
relA	106.987	rpsB	52.800
MW0465	73.078	fab	35.833
rpoC	49.318	pnpA	21.351
murE	24.125	murC	19.000
MW0746	18.636	relA	14.094
MW0413	17.364	MW0465	10.882
fabZ	17.100	rpmB	10.833
ksgA	15.667	leuS	0.098
MW0766	14.283	MW2178	0.095
MW0014	11.830	MW1502	0.081
MW0896	11.233	mvaS	0.080
menD	0.077	bfmBAB	0.076
MW0408	0.076	MW1548	0.069
MW0972	0.075	gyrB	0.067
MW2178	0.070	recU	0.065
mvaS	0.069	menD	0.063
infB	0.065	dnaA	0.062
bfmBAB	0.057	infB	0.043
MW2509	0.055	MW0986	0.037
recU	0.037	codY	0.027
MW0718	0.028	MW1567	0.025
dnaA	0.026	rplS	0.022
MW1567	0.025	tagA	0.022
codY	0.025	tagD	0.021
gyrB	0.022	fni	0.007
tagA	0.018	lexA	0.004
rplS	0.015		
tagD	0.010		
fni	0.010		
lexA	0.008		
MW1548	0.005		

H. How to run all custom R and Python scripts for analysis of Tn-Seq data

The Python scripts for creating igv formatted files, finding the number of TA sites hit, performing the Mann-Whitney U analysis, and performing the Benjamini Hochberg correction were modified from ones made available upon request by Eric Rubin (Harvard Medical School). I wrote the rest. The simulation-based sampling normalization method was adapted from the EL-ARTIST Matlab scripts.

Python scripts for identifying significant differences in mapped-reads

Script: igv_staph_saouhsc_all_11-12.py

This script takes the tabular hopcount files downloaded from Galaxy, and it converts them into igv-formatted files.

Run: python igv_staph_saouhsc_all_11-12.py staph_TA.txt staph_genes.txt hopcountfile.tabular > output.igv

Staph_TA.txt is any tab delimited file listing the genome, start of the TA site, and end of the TA site in three columns.

Staph_genes.txt is any tab delimited file listing each gene, its starting location, and its ending location in three columns.

Hopcountfile.tabular is the tab-delimited hop counts file downloaded from Galaxy

Output will be in igv format

```
1.  import sys
2.
3.  ## Gather all TA sites into dictionary
4.  TAsites = {}
5.  for line in open(sys.argv[1]):
6.      split = line.split()
7.      if split[0] == 'TA':
8.          TAsites[int(split[4])] = [0]
9.  #print TAsites
10.
11. ## Add gene name as second string in dictionary value
```

```
12. for line in open(sys.argv[2]):
13.     split = line.split('\t')
14.     name = split[0]; start = int(split[1]); end = int(split[2])+1
15.     for i in range(start, end):
16.         if i in TAsites: TAsites[i].append(name)
17. #print TAsites
18.
19. ## Add read counts to first item in dictionary value list
20. whole_text=file(sys.argv[3]).read()
21. arrayed = whole_text.split('\n')
22. #lenarrayed = len(arrayed)
23. #print lenarrayed
24. #    print line
25. for line in arrayed:
26.     split = line.split("\t")
27.     #print split[1]
28.     #split = line.split('\t')
29.     pos = 0 ## set pos as int denoting position
30.     if len(split)>= 6 and split[1].isdigit() == True:
31.         if int(split[4]) != 0:
32.             pos = int(split[1]) ##
33.             if pos+15 in TAsites:
34.                 TAsites[pos+15][0] = TAsites[pos+15][0]+int(split[4])
35.             elif pos+14 in TAsites:
36.                 TAsites[pos+14][0] = TAsites[pos+14][0]+int(split[4])
37.         if int(split[5]) != 0:
38.             pos = int(split[1])
39.             if pos-15 in TAsites:
40.                 TAsites[pos-15][0] = TAsites[pos-15][0]+int(split[5])
41.             elif pos-16 in TAsites:
42.                 TAsites[pos-16][0] = TAsites[pos-16][0]+int(split[5])
43.         else:
44.             pass
45. ## Print
46.
47. keys = TAsites.keys()
48. keys.sort()
49.
50. #print "#track name=vanc10_all color=255,255,255 altColor=RRR,GGG,BBB maxHeightPixels=1
    28:128:11 graphType=bar midRange=20:80 midColor=200,200,200 viewLimits=0:600000 windowi
    ngFunction=maximum coords=0 scaleType=linear featureVisivilityWindow=50 gffTags=off"
51. for k in keys:
52.     print 'SAOHSC', '\t', int(k), '\t', int(k) + 2, '\t', TAsites[k][0],
53.     if len(TAsites[k]) > 1: print '\t', TAsites[k][1]
54.     else: print
```

Script: Hit_Sites_Counter.py

This script only takes one argument, the igv file created above. It outputs the total possible

number of TA sites, the number of reads that map to TA sites, and the number of TA sites those

reads map to.

Run: python Hit_Sites_Counter.py output.igv

```
1.  import sys
2.
3.  TA = 0; Reads = 0; HitSites = 0
4.
5.  for line in open(sys.argv[1]):
6.      split=line.split()
7.      if len(split) > 3:
8.          if split[3][0].isdigit() == True:
9.              TA += 1
10.             Reads += float(split[3])
11.             if float(split[3]) > 0: HitSites += 1
12.
13. print 'TA = %d' % (TA)
14. print 'Reads = %d' % (Reads)
15. print 'Sites Hit = %d' % (HitSites)
```

Script: combineigv.py

If multiple biological replicates are done or if you simply want to combine multiple igv files, this script will do it. Takes any number of igv files as arguments, outputs combined igv file.

Run: python igv1.igv, igv2.igv, igv3.igv, etc. > combinedigv.igv

```
1.  import sys
2.
3.  #need to make dictionary of TA site and reads
4.
5.  TAsites = {}
6.  for line in open(sys.argv[1]):
7.      split = line.split('\t')
8.      TAsites[int(split[1])]=[]
9.
10. #print TAsites
11.
12. for arg in sys.argv[1:]:
13.     for line in open(arg):
14.         split = line.split('\t')
15.         site = int(split[1])
16.         reads = int(split[3].rstrip())
17.         if site in TAsites:
18.             TAsites[site].append(reads)
19.
20. #print TAsites
21.
22. #Next add together read numbers and output as original igv
23. allreads = {}
24. for k,v in TAsites.iteritems():
25.     sumreads = sum(v)
26.     allreads[k]=sumreads
27. #print allreads
28.
29. for line in open(sys.argv[1]):
30.     split = line.split('\t')
31.     #print len(split)
```

```
32.      site = int(split[1])
33.      if site in allreads:
34.          if len(split)<5:
35.              print "%s\t%d\t%d\t%d" % (split[0],site,int(split[2]),allreads[site])
36.          elif len(split)<6:
37.              print "%s\t%d\t%d\t%d\t%s" % (split[0],site,int(split[2]),allreads[site],sp
    lit[4].rstrip())
```

Script: normalization.py

This script is essentially a translation of the EL-ARTIST normalization method from MatLab to Python. It takes two arguments. The first should be the igv file with more TA sites hit, and the second the one with fewer TA sites hit. It outputs the first argument normalized down to the second argument's level of diversity.

Run: python normalization.py moretasiteshit.igv fewertasiteshit.igv > normalizedfirstigvfile.igv

```
1.  from math import *
2.  from numpy import *
3.
4.  TActrl = 0; Readsctrl = 0; HitSitesctrl = 0
5.  inputreads = []
6.  saouhsc= []
7.  start = []
8.  end = []
9.  window = []
10. for line in open(sys.argv[1]):
11.     split=line.split('\t')
12.     if len(split) > 4:
13.         if split[3][0].isdigit() == True:
14.             inputreads.append(float(split[3]))
15.             saouhsc.append(split[0])
16.             start.append(split[1])
17.             end.append(split[2])
18.             window.append(split[4].rstrip('\n'))
19.             TActrl += 1
20.             Readsctrl += float(split[3])
21.             if float(split[3]) > 0: HitSitesctrl += 1
22.
23. TAexp = 0; Readsexp = 0; HitSitesexp = 0
24. for line in open(sys.argv[2]):
25.     split=line.split()
26.     if len(split) > 3:
27.         if split[3][0].isdigit() == True:
28.             TAexp += 1
29.             Readsexp += float(split[3])
30.             if float(split[3]) > 0: HitSitesexp += 1
31.
32. TAproportion =  float(HitSitesexp)/HitSitesctrl
33.
34. inputproportion = []
```

```
35. for i in inputreads:
36.     inputproportion.append(float(i)/Readsctrl)
37. inputproportiontanorm = []
38. for i in inputproportion:
39.     inputproportiontanorm.append(float(i)*TAproportion)
40.
41. inputproportiontanorm.append(TAproportion-1)
42.
43. multinominputsample = numpy.random.multinomial(Readsexp,inputproportiontanorm,100)
44. multinominputsample = multinominputsample.T
45. for i in range(len(multinominputsample[0])):
46.     summulti = sum(multinominputsample[:,i])
47. multisum = numpy.repeat(summulti,100)
48. difference=numpy.repeat(multisum-multinominputsample[-1,0],100)
49. correctionfactor = round(float(Readsexp)/difference[0],4)
50. correctedinput = [i*correctionfactor for i in multinominputsample]
51. bootstrapcontrol = numpy.delete(correctedinput, (-1), axis=0)
52. avgbootstrapcontrol = bootstrapcontrol.mean(axis=1)
53.
54. x = 0
55. for x in range(0,len(avgbootstrapcontrol)):
56.     print saouhsc[x], '\t', start[x], '\t', end[x], '\t', avgbootstrapcontrol[x], '\t',
     window[x]
57.     x+=1
```

Script: mannwhitneyu_8-24-15.py

This program collects the reads mapping to each gene and identifies statistically significant differences between the control and experimental condition using the Mann-Whitney U test. It also performs some other calculations on the data which can be used to further analyze the data.

Run: python mannwhitneyu_8-24-15.py staph_genes.txt control.igv experiment.igv > output.txt

Staph_genes.txt is tab-delimited staph genes file described above

Control.igv is igv file from control datasets

Experiment.igv is igv file from experiment datasets

Output.txt is a tab-delimited file that can be opened in Microsoft Excel. Column 1: gene loci, column 2, number of TA sites in that gene, column 3: number reads mapping to that gene in the control data, column 4: number of reads mapping to that gene in the experimental data, column 5: U statistic, column 6: p-value, column 7: ratio in reads in the experiment normalized to the control, 8: length of the gene, column 9: reads mapping to that gene in the control normalized to

243

gene length, column 10: same as column 9 but for the experimental condition, column 11:

location of the start of the gene, column 12: location of the end of the gene.

```
1.  import sys, random, scipy, numpy
2.  from math import *
3.  from scipy import stats
4.
5.
6.  ## DEFINE GENE NAMES, STARTS, ENDS
7.
8.  genes = {}; genereads = {}
9.
10. init = 0
11. count = 1
12.
13. for line in open(sys.argv[1]):
14.     split = line.split('\t')
15.     name = split[0]; start = int(split[1]); end = int(split[2])
16.     genes[name] = [start, end]
17.     genereads[name] = [[],[]]
18. #print genes
19.
20. genelist = genes.keys()
21. genelist.sort()
22. #print genelist
23. #print genereads
24.
25. ## DEFINE RATIO BETWEEN TOTAL READ COUNTS AS CORRECTION
26. Lib1Reads = 0; Lib2Reads = 0
27. Lib1TA = 0; Lib2TA = 0
28.
29. for line in open(sys.argv[2]):
30.     #print line
31.     split = line.split()
32.     #print split[3][0]
33.     if len(split)>1 and split[3][0].isdigit()==True:
34.         Lib1TA += 1
35.         Lib1Reads += float(split[3])
36. #print Lib1Reads
37.
38. for line in open(sys.argv[3]):
39.     split = line.split()
40.     if len(split)>1 and split[3][0].isdigit()==True:
41.         Lib2TA += 1
42.         Lib2Reads += float(split[3])
43.
44. ratio = float(Lib1Reads)/float(Lib2Reads)
45.
46. ## SET DICTIONARIES WITH GENE READS
47.
48. ## set first for Library 1
49. for line in open(sys.argv[2]):
50.     split = line.split()
51.     #print split
52.     if len(split) > 4 and split[3][0].isdigit() == True and split[4] in genereads:
53.         site = int(split[1]); start = int(genes[split[4]][0]); end = int(genes[split[4]
    ][1])
```

```
54.         length = end - start
55.         if (start+(0.03*length)) <= site <= (end-(0.03*length)):
56.             genereads[split[4]][0].append(float(split[3]))
57.     #else:
58.     # print split [4]
59.
60. ## set for Library 2
61. for line in open(sys.argv[3]):
62.     split = line.split()
63.     if len(split) > 4 and split[3][0].isdigit() == True and split[4] in genereads:
64.         site = int(split[1]); start = int(genes[split[4]][0]); end = int(genes[split[4]
    ][1])
65.         length = end - start
66.         if (start+(0.03*length)) <= site <= (end-(0.03*length)):
67.             genereads[split[4]][1].append(float(split[3]))
68.
69. ## MANN-WHITNEY U TEST FOR ALL GENES
70. #print genereads
71. #print len(genelist)
72. for i in range(len(genelist)):
73.     #print genelist[i]
74.     SAO = genelist[i]
75.     start = genes[SAO][0]; end = genes[SAO][1]
76.     #print end
77.     Lib1Counts = genereads[SAO][0]
78.     Lib2Counts = genereads[SAO][1]
79.     TA = len(Lib1Counts)
80.     readslib1 = sum(Lib1Counts)
81.     readslib2 = sum(Lib2Counts)
82.     length = end - start
83.     if length > 0:
84.         index1 = readslib1/length
85.         index2 = readslib2/length
86.     if TA == 0:
87.         print "%s\t%d\t%s" % (SAO,TA,'Gene Has No TAs')
88.     else:
89.         if sum(Lib1Counts) == 0 and sum(Lib2Counts) ==0:
90.             print "%s\t%d\t%s\t%s\t%s\t%s\t%s\t%d\t%s\t%s\t%r\t%r" % (SAO,TA,'0','0','0
    ','0','0',length,'0','0',start,end)
91.         else:
92.             try:
93.                 U, p_val = scipy.stats.mannwhitneyu(Lib1Counts,Lib2Counts)
94.                 CountRatio = float(sum(Lib2Counts)+1)/float(sum(Lib1Counts)+1)
95.                 #print "%s\t%d\t%d\t%0.5f\t%0.3f" % (SAO,TA,U,p_val,CountRatio)
96.                 print "%s\t%d\t%d\t%d\t%d\t%0.100f\t%0.10f\t%d\t%0.10f\t%0.10f\t%r\t%r"
    %
97. (SAO,TA,readslib1,readslib2,U,p_val,CountRatio,length,index1,index2,start,end)
98.             except (ValueError):
99.                 pass
```

Script: gene_fdr_taboutput.py

This script corrects for multiple hypothesis testing by modifying the p-value using the Benjamini

Hochberg procedure. It outputs a tab-delimited file with two columns: the gene locus and the

new p-value. This can be manually inserted into the Mann-Whitney output file. Note: This script does not work if there are non-numbers in your Mann-Whitney p-value column. Sometimes, SAOUHSC_01447 results in a p-value = nan. If script does not work, this could be why.

Run: python gene_fdr_taboutput.py mannwhitneyoutput.txt > corrpvalue.txt

Mannwhitneyoutput.txt should be the output of the previous script. Adding headings to columns will mess it up. After performing this analysis, headings can be added.

```
1.  import operator, random, sys
2.
3.  ## Create Dictionary with genes and P-vals {gene:[p-val],...}
4.
5.  gene_dict = {}
6.
7.  for line in open(sys.argv[1]):
8.      split = line.split('\t')
9.      #print split
10.     if len(split) >= 7:
11.         gene_dict[split[0]] = [split[5]]
12.
13. #print gene_dict
14. ## Function that calculates Benjamini-Hochberg FDR q-
    values. Argument required is ordered list of p-values.
15.
16. def bh_qvalues(pv):
17.     """
18.     Return Benjamini-Hochberg FDR q-values corresponding to p-values C{pv}.
19.
20.     This function implements an algorithm equivalent to L{bh_rejected} but
21.     yields a list of 'adjusted p-values', allowing for rejection decisions
22.     based on any given threshold.
23.
24.     @type pv: list
25.     @param pv: p-values from a multiple statistical test
26.
27.     @rtype: list
28.     @return: adjusted p-values to be compared directly with the desired FDR
29.       level
30.     """
31.     if not pv:
32.         return []
33.     m = len(pv)
34.     args, pv = zip(*sorted(enumerate(pv), None, operator.itemgetter(1)))
35.     #print args
36.     qvalues = m * [0]
37.     mincoeff = pv[-1]
38.     qvalues[args[-1]] = mincoeff
39.     for j in xrange(m-2, -1, -1):
40.         coeff = m*pv[j]/float(j+1)
41.         if coeff < mincoeff:
42.             mincoeff = coeff
43.         qvalues[args[j]] = mincoeff
44.     return qvalues
```

```
45.
46. ## Create ordered list of p_values with their corresponding genes
47.
48. genes_list = []
49.
50. for k,v in gene_dict.iteritems():
51.     genes_list.append([v[0],k])
52.
53. genes_list.sort()
54.
55. p_vals = []; genes = []
56.
57. for i in genes_list:
58.     p_vals.append(float(i[0]))
59.     genes.append(i[1])
60. #print p_vals
61.
62. ## Assign q values for each p value
63.
64. q_vals = bh_qvalues(p_vals)
65. #print q_vals
66.
67. ## Return q values to dictionary
68.
69. for i in range(len(p_vals)):
70.     if genes[i] in gene_dict:
71.         gene_dict[genes[i]].append(q_vals[i])
72. #print gene_dict
73. ## Print
74.
75. for line in open(sys.argv[1]):
76.     split = line.split('\t')
77.     print "%s\t" % split[0],
78.     if split[0] in gene_dict: print gene_dict[split[0]][1],
79.     print
```

Python scripts for identifying upregulated genes

Script: igv_staph_saouhsc_plus.py and igv_staph_saouhsc_minus.py

These scripts are both converts from tabular hopcount files to igv-formatted files and are run in

the same way as igv_staph_saouhsc.py. The only differences are that the output for one

consists of only reads mapping to the plus strand and the other consists of reads mapping only

to the minus strand

Run: python igv_staph_saouhsc_plus.py staph_TA.txt windowstable.txt hopcount.tabular >

outputplusstrand.igv

Run: python igv_staph_saouhsc_minus.py staph_TA.txt staph_genes.txt hopcount.tabular >

outputminusstrand.igv

Windows_table.txt is a tab-delimited file that has split the entire genome into 270bp segments or

windows. It consists of column 1: genome name, column 2: start of the window, and column 3:

end of the window

Igv_staph_saouhsc_plus.py

```
1.  import sys
2.
3.  ## Gather all TA sites into dictionary
4.
5.  TAsites = {}
6.
7.  for line in open(sys.argv[1]):
8.      split = line.split()
9.      if split[0] == 'TA':
10.         TAsites[int(split[4])] = [0]
11. #print TAsites
12. ## Add gene name as second string in dictionary value
13.
14. for line in open(sys.argv[2]):
15.     split = line.split('\t')
16.     name = split[0].rstrip(); start = int(split[1]); end = int(split[2])+1
17.     for i in range(start, end):
18.         if i in TAsites: TAsites[i].append(name)
19. #print TAsites
20.
21.
22.
23. ## Add read counts to first item in dictionary value list
24. whole_text=file(sys.argv[3]).read()
25. arrayed = whole_text.split('\n')
26. #lenarrayed = len(arrayed)
27. #print lenarrayed
28. #    print line
29. for line in arrayed:
30.     split = line.split("\t")
31.     #print split[1]
32.     #split = line.split('\t')
33.     pos = 0 ## set pos as int denoting position
34.     if len(split)>= 6 and split[1].isdigit() == True:
35.         if int(split[4]) != 0:
36.             pos = int(split[1]) ##
37.             if pos+15 in TAsites:
38.                 TAsites[pos+15][0] = TAsites[pos+15][0]+int(split[4])
39.             elif pos+14 in TAsites:
40.                 TAsites[pos+14][0] = TAsites[pos+14][0]+int(split[4])
41.             else:
42.                 pass
43. ## Print
44.
45. keys = TAsites.keys()
```

248

```
46. keys.sort()
47.
48. for k in keys:
49.     print 'SAOHSC', '\t', int(k), '\t', int(k) + 2, '\t', TAsites[k][0],
50.     if len(TAsites[k]) > 1: print '\t', "'""'",TAsites[k][1],"'"
51.     else: print
52. import sys
```

igv_staph_saouhsc_minus.py

```
1.  import sys
2.
3.  ## Gather all TA sites into dictionary
4.
5.  TAsites = {}
6.
7.  for line in open(sys.argv[1]):
8.      split = line.split()
9.      if split[0] == 'TA':
10.         TAsites[int(split[4])] = [0]
11. #print TAsites
12. ## Add gene name as second string in dictionary value
13. for line in open(sys.argv[2]):
14.     split = line.split('\t')
15.     name = split[0].rstrip(); start = int(split[1]); end = int(split[2])+1
16.     for i in range(start, end):
17.         if i in TAsites: TAsites[i].append(name)
18. #print TAsites
19.
20.
21.
22. ## Add read counts to first item in dictionary value list
23.
24.
25. whole_text=file(sys.argv[3]).read()
26. arrayed = whole_text.split('\n')
27. #lenarrayed = len(arrayed)
28. #print lenarrayed
29. #    print line
30. for line in arrayed:
31.     split = line.split("\t")
32.     #print split[1]
33.     #split = line.split('\t')
34.     pos = 0 ## set pos as int denoting position
35.     if len(split)>= 6 and split[1].isdigit() == True:
36.         if int(split[5]) != 0:
37.             pos = int(split[1])
38.             if pos-15 in TAsites:
39.                 TAsites[pos-15][0] = TAsites[pos-15][0]+int(split[5])
40.             elif pos-16 in TAsites:
41.                 TAsites[pos-16][0] = TAsites[pos-16][0]+int(split[5])
42.             else:
43.                 pass
44. ## Print
45.
46. keys = TAsites.keys()
47. keys.sort()
48.
```

```
49. for k in keys:
50.     print 'SAOHSC', '\t', int(k), '\t', int(k) + 2, '\t', TAsites[k][0],
51.     if len(TAsites[k]) > 1: print '\t', "'""'",TAsites[k][1],"'"
52.     else: print
```

Script: normalization.py

Essentially the same as the previously-described normalization file.

Run: python normalization.py moretasiteshit.igv fewertasiteshit.igv > normalizedfirstigvfile.igv

```
1.  import sys, numpy
2.  from math import *
3.  from numpy import *
4.
5.  TActrl = 0; Readsctrl = 0; HitSitesctrl = 0
6.  inputreads = []
7.  saouhsc= []
8.  start = []
9.  end = []
10. window = []
11. for line in open(sys.argv[1]):
12.     split=line.split('\t')
13.     if len(split) > 3:
14.         if split[3][0].isdigit() == True:
15.             inputreads.append(float(split[3]))
16.             saouhsc.append(split[0])
17.             start.append(split[1])
18.             end.append(split[2])
19.             window.append(split[4].rstrip('\n'))
20.             TActrl += 1
21.             Readsctrl += float(split[3])
22.             if float(split[3]) > 0: HitSitesctrl += 1
23.
24. TAexp = 0; Readsexp = 0; HitSitesexp = 0
25. for line in open(sys.argv[2]):
26.     split=line.split()
27.     if len(split) > 3:
28.         if split[3][0].isdigit() == True:
29.             TAexp += 1
30.             Readsexp += float(split[3])
31.             if float(split[3]) > 0: HitSitesexp += 1
32.
33. TAproportion =  float(HitSitesexp)/HitSitesctrl
34.
35. inputproportion = []
36. for i in inputreads:
37.     inputproportion.append(float(i)/Readsctrl)
38. inputproportiontanorm = []
39. for i in inputproportion:
40.     inputproportiontanorm.append(float(i)*TAproportion)
41.
42. inputproportiontanorm.append(TAproportion-1)
43.
44. multinominputsample = numpy.random.multinomial(Readsexp,inputproportiontanorm,100)
45. multinominputsample = multinominputsample.T
46. for i in range(len(multinominputsample[0])):
```

```
47.       summulti = sum(multinominputsample[:,i])
48. multisum = numpy.repeat(summulti,100)
49. difference=numpy.repeat(multisum-multinominputsample[-1,0],100)
50. correctionfactor = round(float(Readsexp)/difference[0],4)
51. correctedinput = [i*correctionfactor for i in multinominputsample]
52. bootstrapcontrol = numpy.delete(correctedinput, (-1), axis=0)
53. avgbootstrapcontrol = bootstrapcontrol.mean(axis=1)
54.
55. x = 0
56. for x in range(0,len(avgbootstrapcontrol)):
57.     print saouhsc[x], '\t', start[x], '\t', end[x], '\t', avgbootstrapcontrol[x], '\t',
    window[x]
58.     x+=1
```

Script: promoter_analysis.py

This takes the igv files and the normalized igv files and identifies TA sites with a large increase in number of reads compared to the control and a preference for insertion orientation. If experiment condition hit more TA sites and is the normalized one, arguments are run in the same order.

Run: python promoter_analysis.py normalizedctrlplusreads.igv experimentplusreads.igv normalizedminusreads.igv experimentminusreads.igv > output.txt

Output is a tab-delimited file with column 1: window name, column 2: TA site location, column 3: hit direction (if plus, upregulating downstream genes, if minus upregulating upstream genes, column 4: log(ratio) plus strand, column 5: log(ratio)minus strand, column 6: log(reads) plus strand, and column 7: log(reads) minus strand. Values are converted to log(1+value) to make distribution more normal.

```
1.  import sys, numpy
2.  from math import *
3.  from numpy import *
4.
5.  TAdict = {}
6.  hitdict = {}
7.  #extract TAsite location and put it in as key for dictionary
8.  for line in open(sys.argv[1]):
9.      split = line.split('\t')
10.     TAsite = int(split[1])
11.     TAdict[TAsite] = []
12.     hitdict[TAsite] = []
13.
```

```
14. #put reads into dictionary
15. for line in open(sys.argv[1]):
16.     split = line.split('\t')
17.     name = split[4].rstrip('\n')
18.     TAdict[int(split[1])].append(name)
19.     reads = int(round(float(split[3]),0))
20.     reads = reads + 1
21.     TAdict[int(split[1])].append(reads)
22.
23. for line in open(sys.argv[2]):
24.     split = line.split('\t')
25.     reads = int(round(float(split[3]),0))
26.     reads = reads+1
27.     TAdict[int(split[1])].append(reads)
28.
29. for line in open(sys.argv[3]):
30.     split = line.split('\t')
31.     reads = int(round(float(split[3]),0))
32.     reads = reads+1
33.     TAdict[int(split[1])].append(reads)
34.
35. for line in open(sys.argv[4]):
36.     split = line.split('\t')
37.     reads = int(round(float(split[3]),0))
38.     reads = reads+1
39.     TAdict[int(split[1])].append(reads)
40.
41. #print TAdict
42. #calculate and add ratios to dictionary
43. for k in TAdict:
44.     #print TAdict[k][3]
45.     ratioplus = float(TAdict[k][2])/TAdict[k][1]
46.     #print ratioplus
47.     TAdict[k].append(ratioplus)
48.     logratioplus = log10(ratioplus)
49.     #print logratioplus
50.     TAdict[k].append(logratioplus)
51.     ratiominus = float(TAdict[k][4])/TAdict[k][3]
52.     #print ratiominus
53.     TAdict[k].append(ratiominus)
54.     logratiominus = log10(ratiominus)
55.     #print logratiominus
56.     TAdict[k].append(logratiominus)
57.     readsplus = int(TAdict[k][2])
58.     logreadsplus = log10(readsplus)
59.     TAdict[k].append(logreadsplus)
60.     logreadsminus = log10(TAdict[k][4])
61.     TAdict[k].append(logreadsminus)
62.
63. #print TAdict
64. #make lists of logplus and log minus also of reads for exp
65. logplus = []
66. logminus = []
67. logreadsplus = []
68. logreadsminus = []
69. for k in TAdict:
70.     logplus.append(TAdict[k][6])
71.     logminus.append(TAdict[k][8])
72.     logreadsplus.append(TAdict[k][9])
73.     logreadsminus.append(TAdict[k][10])
```

```
74.
75. #calculate average and standard deviations - ADJUST NUMBERS HERE!
76. meanplus = float(sum(logplus))/len(logplus)
77. meanminus = float(sum(logminus))/len(logminus)
78. meanreadsplus = float(sum(logreadsplus))/len(logreadsplus)
79. meanreadsminus = float(sum(logreadsminus))/len(logreadsminus)
80. twosdplus = 7*(numpy.std(logplus, axis=0)) #This value can be modified to increase/decr
    ease cutoffs
81. twosdminus = 7*(numpy.std(logminus, axis=0)) #This value can be modified to increase/de
    crease cutoffs
82. sdreadsplus = 7*(numpy.std(logreadsplus, axis=0)) #This value can be modified to increa
    se/decrease cutoffs
83. sdreadsminus = 7*(numpy.std(logreadsminus, axis=0)) #This value can be modified to incr
    ease/decrease cutoffs
84. ratiocutoffplus = meanplus+twosdplus
85. ratiocutoffminus = meanminus+twosdminus
86. readscutoffplus = meanreadsplus+sdreadsplus
87. readscutoffminus = meanreadsminus+sdreadsminus
88.
89. #filter and print
90. keys = TAdict.keys()
91. keys.sort()
92.
93. hitsplus = "hitsplus"
94. hitsminus = "hitsminus"
95. #find hits
96. for k in keys:
97.     if TAdict[k][6] > ratiocutoffplus and TAdict[k][9] > readscutoffplus:
98.         TAdict[k].append(hitsplus)
99.     elif TAdict[k][8] > ratiocutoffplus and TAdict[k][10] > readscutoffminus:
100.            TAdict[k].append(hitsminus)
101.        #print TAdict
102.        for k in keys:
103.            #print TAdict[k][-1]
104.            #print isinstance(TAdict[k][-1], str)
105.            if isinstance(TAdict[k][-1],str)==True:
106.                #print TAdict[k][-1]
107.                #print k
108.                hitdict[k].append(TAdict[k][0])
109.                hitdict[k].append(TAdict[k][6])
110.                hitdict[k].append(TAdict[k][8])
111.                hitdict[k].append(TAdict[k][9])
112.                hitdict[k].append(TAdict[k][10])
113.                hitdict[k].append(TAdict[k][-1])
114.
115.        hitkeys = hitdict.keys()
116.        hitkeys.sort()
117.
118.        #this is the new printing part
119.        print 'window name \t TAsite \t hitdirection \t logratioplus \t logratiominus \t
    logreadsplus \t logreadsminus'
120.        for k in keys:
121.            if len(hitdict[k]) > 4:
122.                print hitdict[k][0], '\t', k, '\t', hitdict[k][-
    1], '\t', hitdict[k][1], '\t', hitdict[k][2], '\t', hitdict[k][3], '\t', hitdict[k][4]

123.
124.        #print hitdict
```

Script: driver.py

Running the promoter analysis for each sample one script at a time is somewhat time-intensive, so this script can be run on everything at once. All control and experiment tabular files need to be in a directly with only the other scripts necessary for running this analysis. In addition, files must be named in a certain way. Control files must be named "control_tnconstruct.tabular" and experiment files must be named "experimentalcondition_tnconstruct.tabular" where tnconstruct = blunt, cap, dual, erm, pen or tuf.

Run: python driver.py

This file automatically outputs the files in the same format described above.

```
1.  import os
2.  from subprocess import call
3.
4.  def without_extension(filename):
5.      return os.path.splitext(filename)[0]
6.
7.  def get_hit_sites(path):
8.      TA = 0; Reads = 0; HitSites = 0
9.
10.     input_file = open(path)
11.     for line in input_file:
12.         split = line.split()
13.         if len(split) > 3:
14.             if split[3][0].isdigit() == True:
15.                 TA += 1
16.                 Reads += float(split[3])
17.                 if float(split[3]) > 0: HitSites += 1
18.     input_file.close()
19.     return HitSites
20.
21. inputs = [path for path in os.listdir('.') if path.endswith('.tabular')]
22. igv_files = []
23. for data in inputs:
24.     name = without_extension(data)
25.     map_to_ta_plus = name + 'plus.igv'
26.     map_to_ta_minus = name + 'minus.igv'
27.     run_command = "python %s Staph_TA.txt windows_table.txt " + data
28.     os.system(run_command % "igv_staph_saouhsc_plus.py" + " > " + map_to_ta_plus)
29.     os.system(run_command % "igv_staph_saouhsc_minus.py" + " > " + map_to_ta_minus)
30.     igv_files.append(map_to_ta_plus)
31.     igv_files.append(map_to_ta_minus)
32.
33.
34. def get_promoter_analysis_name(filename):
35.     name = without_extension(filename)
36.     # names are usually like moe16_tufminus
37.     return name.split('_')[1]
```

254

```python
38.
39. # Normalize!
40. # We need to group the igv files by promoter analysis name so we can
41. # pass the control and experiment to normalize
42. group_labels = set(map(get_promoter_analysis_name, igv_files))
43. grouped_items = []
44. for group_label in group_labels:
45.     group = []
46.     for igv_file in igv_files:
47.         if group_label in igv_file: group.append((igv_file, get_hit_sites(igv_file)))
48.     grouped_items.append(group)
49.
50. # We want to sort by the hit sit count. We already stored this count in the tuple:
51. # [[(ctrl_path1, hit_site), (expr_path1, hit_site)], [(ctrl_path2, hit_site), (expr_path2, hit_site)]]
52. normalized_args = [sorted(group, key=lambda item: item[1], reverse=True) for group in grouped_items]
53.
54. # Extract only path from each tuple in the groups
55. paths = [(group[0][0], group[1][0]) for group in normalized_args]
56.
57. # We will generate a list of the inputs for each pair to promoter_analysis
58. # will have to repair plus and minus later
59. # TODO make this not suck
60. promoter_analysis_inputs = []
61. for arg1, arg2 in paths:
62.     # output arg1 with norm before extension
63.     destination = without_extension(arg1) + ".norm.igv"
64.     os.system("python normalization.py %s %s > %s" % (arg1, arg2, destination))
65.     if arg1.startswith('control_'):
66.         promoter_analysis_inputs.append((destination, arg2))
67.     else:
68.         promoter_analysis_inputs.append((arg2, destination))
69.
70. promoter_args = []
71. # We need to re-pair plus with minus before calling promoter analysis
72. plus_strands = [strand for strand in promoter_analysis_inputs if "plus" in strand[0]]
73. minus_strands = [strand for strand in promoter_analysis_inputs if "plus" not in strand[0]]
74. for plus_strand in plus_strands:
75.     promoter_name = str.replace(get_promoter_analysis_name(str.replace(plus_strand[0], '.norm', '')), 'plus', '')
76.     for minus_strand in minus_strands:
77.         if promoter_name in minus_strand[0]:
78.             promoter_args.append(plus_strand + minus_strand)
79.
80. for arglist in promoter_args:
81.     print arglist
82.     # Changed line!
83.     output_name = str.replace(str.replace(arglist[0].split('_')[1], '.norm.igv', ''), 'plus', '') + ".promoter.out"
84.     command = "python promoteranalysis.py %s %s %s %s > " + output_name
85.     # Changed line!
86.     os.system(command % (arglist[1],arglist[0],arglist[3],arglist[2]))
```

Python scripts for identifying genes with similar resistance and sensitization patterns

Script: mwu_to_cluster_11-1-15.py

This converts a mann-whitney U output file and a false discovery rate file to put the data into the format required for performing the k-nearest neighbors classifier analysis. It modifies the ratio to account for a few number of reads in both the control and the experiment, and it also calculates the normalized fitness value which we use in these clustering analyses.

Run: python mwu_to_cluster.py mwu_sample.txt fdr_sample.txt > outputforclustering.txt

Output consists of a csv (comma-separated) file with column1: gene locus, column 2: number of TA sites, column 3: length, column 4: number of reads in the control sample mapping to this gene, column 5: number of reads in the experimental sample mapping to this gene, column 6: raw ratio value, column 7: corrected p-value, column 8: modified ratio, column 9: Reads mapping to the gene in the experiment normalized to the length of the gene, column 10: normalized fitness value (used for clustering)

```python
1.  import operator, random, sys
2.
3.  ## Create Dictionary with genes and P-vals {gene:[p-val],...}
4.
5.  gene_dict = {}
6.
7.  #find total reads in ctrl and exp
8.  readcountctrl = 0
9.  readcountexp = 0
10. for line in open(sys.argv[1]):
11.     split = line.split('\t')
12.     #print split
13.     if split[0][0:3]=='SAO':
14.         try:
15.             readcountctrl = readcountctrl + int(split[2])
16.             readcountexp = readcountexp + int(split[3])
17.         except Exception:
18.             pass
19.
20.
21. mincountctrl = float(readcountctrl)/10000
22. mincountexp = float(readcountexp)/10000
23.
24. #put important info in a dictionary: 0-#TA sites, 1-length, 2-readsctrl,3-readsexp, 4-
        rawratio, 5-modratio, 6-index, 7-rawfitscore
25. for line in open(sys.argv[1]):
26.     split = line.split('\t')
27.     if split[0][0:3]=='SAO' and int(split[1])>=0:
```

```
28.          try:
29.              readsctrl=int(split[2])
30.          except Exception:
31.              pass
32.          if readsctrl < mincountctrl: readsctrl = mincountctrl
33.          try:
34.              readsexp = int(split[3])
35.          except Exception:
36.              pass
37.          if readsexp <mincountexp: readsexp = mincountexp
38.          ratio = float(readsexp)/readsctrl
39.          try:
40.              TAsites = int(split[1])
41.              length = int(split[7])
42.              rawratio = float(split[6])
43.              #print split[6]
44.              index = float(readsexp)/length
45.              rawfit = ratio*index
46.              gene_dict[split[0]] = [TAsites,length,int(split[2]),int(split[3]),rawratio,
    ratio,index,rawfit]
47.          except Exception:
48.              pass
49.
50. #Add corrected pval from fdr file into dictionary
51. #0-#TA sites, 1-length, 2-readsctrl,3-readsexp, 4-rawratio, 5-modratio, 6-index, 7-
    rawfitscore, 8-corrpval
52. for line in open(sys.argv[2]):
53.     split = line.split('\t')
54.     if split[0] in gene_dict:
55.         gene_dict[split[0]].append(float(split[1].rstrip()))
56.
57. #print gene_dict
58.
59. #Make fitvals a list instead of a dict for calculating norm-fit value
60. fitvals = []
61. for k,v in gene_dict.iteritems():
62.     fitvals.append([v[7],k])
63.
64. fitvals.sort()
65. count=0
66. for f in fitvals:
67.     f.append(count)
68.     count+=1
69.     normorder = float(count)/len(fitvals)
70.     f.append(normorder)
71. #print fitvals
72.
73. #put normalized fitval into genedict
74. for f in fitvals:
75.     gene = f[1]
76.     if gene in gene_dict:
77.         gene_dict[gene].append(f[3])
78. #print gene_dict
79.
80. genes = gene_dict.keys()
81. genes.sort()
82.
83. for g in genes:
84.     #print len(gene_dict[g])
85.     if len(gene_dict[g]) > 6:
```

```
86.        print "%s,%d,%d,%d,%d,%0.10f,%0.100f,%0.10f,%0.10f,%0.6f" % (g,gene_dict[g][0],
    gene_dict[g][1],gene_dict[g][2],gene_dict[g][3],gene_dict[g][4],gene_dict[g][8],gene_di
    ct[g][5],gene_dict[g][6],gene_dict[g][9])
87.
88.        #genename,#TAsites,length,readsctrl,readsexp,rawratio,corrpval,modratio,index,n
    ormfitscore
```

Script: fit_to_array_for_knn_predictgene.py

This script is used to identify the 5 nearest neighbors that have a similar resistance and
sensitization pattern as a gene of interest. Inputs include outputs of mwu_to_cluster.py (as
many as you want). Gene of interest is simply typed in. Essential genes are removed ahead of
time for this analysis.

Run: python fit_to_array_for_knn_predictgene.py essentials.txt clustfile1.txt clustfile2.txt
clustfile3.txt etc SAOUHSC_01234 > output.txt

Essentials.txt is list of essential genes to be removed from the analysis

Output is a list of every gene in the genome and the probability estimates for the list of five most
similar genes. Probability estimates add up to one. Note: probability estimates will be one fewer
than number of genes. You need to move the column down one row after your gene of interest
to identify the correct most similar genes.

```
1.  import sys
2.  import numpy as numpy
3.  from sklearn import neighbors
4.
5.  #put data in array format
6.
7.  y = []
8.  X = []
9.  genenames = {}
10. unknown = []
11. essgenes = []
12.
13. #first value in dictionary is unknown
14. for arg in sys.argv:
15.     if "essentials" in arg:
16.         for line in open(arg):
17.             #print line.rstrip()
18.             essgenes.append(line.rstrip())
19.
20. for arg in sys.argv:
```

```
21.    if arg.endswith('.txt') == True:
22.        for line in open(arg):
23.            split = line.split(',')
24.            if len(split) > 2:
25.                name = split[0]
26.                if any(name in x for x in essgenes):
27.                    pass
28.                else:
29.                    fitval = float(split[9].rstrip())
30.                    if name not in genenames:
31.                        genenames[name]=[fitval]
32.                    else:
33.                        genenames[name].append(fitval)
34.    else:
35.        geneofinterest = arg
36. #print essgenes
37. #print len(genenames)
38. for k,v in genenames.iteritems():
39.    if k == geneofinterest:
40.        unknown.append(v)
41.    else:
42.        X.append(v)
43.        y.append(k)
44. #print X
45.
46. #fit data wiht K nearest neighbors
47. n_neighbors = 5 #this can be increased or decrease
48. h=0.2 #this is the step size in mesh
49.
50. #see http://scikit-
    learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html for diff
    erent options here
51. for weights in ['uniform','distance']:
52.    neigh = neighbors.KNeighborsClassifier(n_neighbors, weights=weights)
53.    neigh.fit(X,y)
54.
55. #predict for unknown
56.
57. for i in neigh.predict_proba(unknown):
58.    listofclosestneighbors = i
59.    for i in listofclosestneighbors:
60.        print i
```

Python scripts for predicting antibiotic mechanism of action using machine learning

Script: fit_to_array_for_knn_2-29.py

This script identifies the antibiotic class most similar to a test/new antibiotic based on the resistance factor pattern for that antibiotic. First, you need to decide which antibiotics to use and how to classify them. Each class receives its own integer value to identify it. The unknown antibiotic must have the word "unknown" in its file name. This version of the script identifies the

unique set of genes in the top 25% and bottom 25% of fitness values for every antibiotic put into

the training set and uses those genes to predict the mechanism of a new antibiotic.

Run: python fit_to_array_for_knn_2-29.py clust_compound1.txt category1 clust_compound2.txt

category2 clust_compound3.txt category3 etc. clust_unknown.txt

Outputs the most similar category as well as the probability estimates for each category.

```python
1.  import sys
2.  import numpy as numpy
3.  from sklearn import neighbors
4.
5.  #put data in array format
6.
7.  y = []
8.  X = []
9.  unknown = []
10.
11. cutoff=0.25
12.
13. for arg in sys.argv:
14.     if arg.endswith('.txt') == True:
15.         if "unknown" in arg:
16.             for line in open(arg):
17.                 split = line.split(',')
18.                 #unknown.append(float(split[9].rstrip()))
19.                 if float(split[5].rstrip()) >= 1-cutoff:
20.                     unknown.append(split[5].rstrip())
21.                 elif float(split[5].rstrip()) <= cutoff:
22.                     unknown.append(split[5].rstrip())
23.                 else:
24.                     pass
25.         else:
26.             fitvals = []
27.             for line in open(arg):
28.                 split = line.split(',')
29.                 #fitvals.append(float(split[9].rstrip()))
30.                 if float(split[5].rstrip()) >= 1-cutoff:
31.                     fitvals.append(split[5].rstrip())
32.                 elif float(split[5].rstrip()) <= cutoff:
33.                     fitvals.append(split[5].rstrip())
34.                 else:
35.                     pass
36.             #print fitvals
37.             X.append(fitvals)
38.     else:
39.         try:
40.             y.append(int(arg))
41.         except ValueError:
42.             pass
43. #print len(fitvals)
44. print len(unknown)
45. #print len(y)
46. #print X
47. #print y
```

```
48.
49. #fit data wiht K nearest neighbors
50.
51. n_neighbors = 2
52. h=0.2 #this is the step size in mesh - can change it to figure out what it does
53. metric = "minkowski"
54.
55. #see http://scikit-
    learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html for diff
    erent options here
56. for weights in ['uniform','distance']:
57.     neigh = neighbors.KNeighborsClassifier(n_neighbors, weights=weights, metric=metric)

58.     neigh.fit(X,y)
59.
60. #predict for unknown
61. print neigh.predict(unknown)
62. print neigh.predict_proba(unknown)
```

R scripts

R scripts can be run all at once or line by line in the RGui. I recommend running them line by line in the Gui to better troubleshoot issues

Script: pca.R

This performs principal component analysis on Tn-Seq data and plots it on a graph. Inputs should be a comma-separated value formatted array containing the treatment conditions and genes you are interested in comparing. It can also be done with every gene in the genome.

```
1.  #Principle components analysis
2.
3.  #set working directory
4.  setwd("C:/Users/etc")
5.
6.  #load data
7.  mydata <- read.csv("data.csv")
8.
9.  # Pricipal Components Analysis
10. # entering raw data and extracting PCs
11. # from the correlation matrix
12. fit <- princomp(mydata, cor=TRUE)
13. summary(fit) # print variance accounted for
14. loadings(fit) # pc loadings
15. plot(fit,type="lines") # scree plot
16. fit$scores # the principal components
17. biplot(fit)
18.
```

```
19. #Plot PCA graph
20. png("pca_donors.png",width=500, height=500, res=100)
21. plot(loadings(fit),type="n", main="Compounds", cex.main=1.5,xlab="Component 1", ylab="C
    omponent 2",cex.lab=1.5)
22. text(loadings(fit),labels=colnames(mydata))
23. dev.off()
```

Script: hierarchicalclustering.R

This script performs hierarchical clustering on any set of Tn-Seq data and outputs a heatmap of the results. Input should be a tab-delimited file containing an array of the antibiotic treatments and genes of interest

```
1.  setwd("C:/Users/etc/")
2.
3.  #load data
4.  data <- read.table("tab-
    delimitedarray.txt",header=T, stringsAsFactors=T, row.names=1)
5.  mydata=na.omit(data)
6.
7.
8.  ## Hierarchical clustering routine
9.  y <- mydata
10. hr <- hclust(as.dist(1-cor(t(y), method="pearson")), method="complete"); hc <-
    hclust(as.dist(1-
    cor(y, method="spearman")), method="complete") # Generates row and column dendrograms.

11. mycl <- cutree(hr, h=max(hr$height)/1.5); mycolhc <-
    rainbow(length(unique(mycl)), start=0.1, end=0.9); mycolhc <-
    mycolhc[as.vector(mycl)] # Cuts the tree and creates color vector for clusters.
12.
13. library(gplots); myheatcol <-
    redgreen(66) # Assign your favorite heatmap color scheme. Some useful examples: colorp
    anel(40, "darkblue", "yellow", "white"); heat.colors(75); cm.colors(75); rainbow(75); r
    edgreen(75); library(RColorBrewer); rev(brewer.pal(9,"Blues")[-
    1]). Type demo.col(20) to see more color schemes.
14.
15. png('graphfitvalknowns1.png', width=720, height=720)
16. heatmap.2(as.matrix(y), Rowv=as.dendrogram(hr), Colv=as.dendrogram(hc), col=myheatcol,
    scale="row", density.info="none", trace="none", RowSideColors=mycolhc, margin=c(20,10),
     cexCol=2.0, cexRow=0.001) # Creates heatmap for entire data set where the obtained clu
    sters are indicated in the color bar.
17. dev.off()
18.
19. x11(height=6, width=2); names(mycolhc) <-
    names(mycl); barplot(rep(10, max(mycl)), col=unique(mycolhc[hr$labels[hr$order]]), hor
    iz=T, names=unique(mycl[hr$order])) # Prints color key for cluster assignments. The num
    bers next to the color boxes correspond to the cluster numbers in 'mycl'.
20.
21. clid <- c(4,6); ysub <- as.matrix(y[names(mycl[mycl%in%clid]),]); hrsub <-
    hclust(as.dist(1-cor(t(ysub), method="pearson")), method="complete")
22.      # Select sub-
    cluster number (here: clid=c(1,2)) and generate corresponding dendrogram.
```

262

```
23.
24. x11(); heatmap.2(ysub, Rowv=as.dendrogram(hrsub), Colv=as.dendrogram(hc), col=myheatcol
    , scale="row", density.info="none", trace="none", RowSideColors=mycolhc[mycl%in%clid])
    # Create heatmap for chosen sub-cluster.
25.
26. #This gets you genes from any point you care about
27. genenames <-
    data.frame(Labels=rev(hrsub$labels[hrsub$order])) # Print out row labels in same order
    as shown in the heatmap.
28. write.table(genenames, file="genenames.txt")
```