Summer 8-15-2015

β-Lactamase Gene Exchange within the Enterobacteriaceae

Mitchell William Pesesky
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Microbiology and Microbial Pathogenesis

Dissertation Examination Committee:
Gautam Dantas, Chair
Carey-Ann Burnham
Barak Cohen
Jeffrey Gordon
Scott Hultgren
David Wang

β-Lactamase Gene Exchange within the *Enterobacteriaceae*

by
Mitchell W Pesesky

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2015
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Abbreviations

HGT – horizontal gene transfer

MRSA – methicillin resistant *Staphylococcus aureus*

HAI – hospital acquired Infection

ESBL – extended spectrum β-lactamase

CRE – carbapenem resistant *Enterobacteriaceae*

PBP – penicillin-binding protein

MDR – multi-drug resistant

MLST – multi-locus sequence typing

ORF – open reading frame

PH – Pakistani hospital samples

WU – Washington University samples

CAI – codon adaptation index

AST – antibiotic susceptibility testing

WGS – whole genome sequencing

RB – rules-based

LR – logistic regression

RF – Random Forest

DHFR – dihydrofolate reductase

QRDR – quinolone resistance determining region

ZOI – zone of inhibition

MIC – minimum inhibitory concentration

Acknowledgments

I received support from the NIGMS Cell and Molecular Biology Training Grant (GM:007067) for my work on each of these projects. These projects also received support from the NIH Director's New Innovator Award (http://commonfund.nih.gov/newinnovator/), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK: http://www.niddk.nih.gov/), and the National Institute of General Medical Sciences (NIGMS: http://www.nigms.nih.gov/), of the National Institutes of Health (NIH) under award numbers DP2DK098089 and R01GM099538 to Gautam Dantas.

I am grateful to my thesis committee for taking the time to give me what always turned out to be spot on guidance.

I would like to thank Siddharth Krishnamurthy for detailed discussions of machine learning algorithms. Thanks to Jeffrey Gordon access to strains used in the mock intestinal community, Matt Hibberd for isolating genomic DNA from those strains, and Jeremiah Faith for providing bioinformatic support. I am very grateful to all members of the Dantas lab for insightful discussions on all of my work. I would also like to thank Tahir Hussain, Erica Pehrsson, Molly Gibson, and Patrick Gonzales in particular for being excellent close collaborators.

To the members of the Brew Crew: thank you for learning a new craft along with me, and for all of our other adventures. It would have been a lot harder to stay sane without you.

I am fortunate to have had the opportunity to have worked with Gautam Dantas, who was the ideal mentor. He gave me independence when I wanted it, advice when I needed it, and optional formality every Friday.

A great deal of credit for everything contained herein belongs with my parents William and Sharon Pesesky, who early on taught me the joys and benefits of a good education, through their words and actions.

Mitchell Pesesky

Washington University in St. Louis

May 2015

Dedicated to my wife and the love of my life, Katie Thacher.

ABSTRACT OF THE DISSERTATION

β-Lactamase Gene Exchange within the *Enterobacteriaceae*

by

Mitchell W Pesesky

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Microbiology and Microbial Pathogenesis

Washington University in St. Louis, 2015

Professor Gautam Dantas, Chair

Antibiotic resistance represents a grave threat to modern medicine's control over infectious disease. Pathogens of the *Enterobacteriaceae* have proven particularly problematic as they can cause a wide variety of infections, and they can be, in some cases, resistant to all antibiotics recommended for use against them. A major part of the threat posed by the *Enterobacteriaceae* is their ability to exchange resistance genes by horizontal gene transfer (HGT). HGT has allowed some *Enterobacteriaceae* to quickly accumulate resistance against diverse antibiotics, and then to spread their resistance gene collection to other pathogenic strains. I explore three aspects of how HGT has affected the spread of resistance in the *Enterobacteriaceae*, with particular focus on the β-lactamase resistance genes: 1) how two, recently discovered β-lactamases have spread between different strains and species in the *Enterobacteriaceae*; 2) the extent to which other phyla in the microbiota can contribute new resistance genes to the *Enterobacteriaceae*; and 3) how whole genome sequencing can be used as a clinical diagnostic to detect antibiotic resistance in the *Enterobacteriaceae*.

There are two major methods by which a resistance gene may increase in prevalence: clonal expansion in a successful strain, or HGT between strains. To determine the contributions of each of these strategies to the success of the KPC and NDM-1 β-lactamases, we use whole genome sequencing and plasmid sequencing to identify the genetic context of the genes encoding each protein. We found that although these genes can sometimes be found in multiple members of the same clade, they are also encoded by much more distant strains and species. Similarly, we identified a few pairs of plasmids with high sequence identity where one carried NDM-1 and the other a KPC, but overall the similarity between any two plasmids was low, even between plasmids carrying the same β-lactamase. This suggests that HGT is playing a large role in the spread of these genes, and that it is mediated not by a single plasmid but by a diverse array of plasmids.

To determine the potential for the *Enterobacteriaceae* to gain antibiotic resistance genes from other members of the microbiota, we applied a parametric measure of HGT to all of the resistance genes from a set of 457 known microbiota strains. Although resistance genes were significantly more likely to have undergone HGT than phylogenetic marker genes, the distance between the two groups was small. We tested the ability of *Escherichia coli*, a member of the *Enterobacteriaceae*, to utilize resistance genes from other phyla using functional metagenomics, pulling genes from a collection of 76 known microbiota strains. We found only a few examples of genes from other phyla that could be used by *E. coli*, again suggesting a minor role for inter-phyla HGT in the antibiotic resistance of the *Enterobacteriaceae*.

One way to counter the spread of antibiotic resistance is through antibiotic stewardship coupled with active monitoring of resistance genes in pathogens. One barrier to this is that the current standard antibiotic resistance diagnostic in hospitals takes two days to determine the

resistance of an isolate, and it does not identify the gene causing resistance for most patients. Whole genome sequencing has been proposed as an alternative method that could return results more quickly, and with gene information. We used a set of 78 *Enterobacteriaceae* clinical isolates to compare two approaches for determining antibiotic susceptibility of a pathogenic isolate from its genome sequence, a rules-based algorithm and a machine-learning algorithm. We found that both algorithms performed similarly, with approximately 90% of predictions matching what was found by *in vitro* phenotyping, but that the machine-learning algorithm showed potential for extension to more specific predictions.

In sum, I explore the scope of how HGT contributes to the increasing prevalence of antibiotic resistance in *Enterobacteriaceae*, and one method that can be used to boost containment efforts.

Chapter 1: Introduction to antibiotic resistance in the *Enterobacteriaceae*

Antibiotics are one of the most powerful tools of modern medicine. By providing a measure of control over bacterial infections, antibiotics save or improve millions of lives each year (1). Antibiotic resistance by bacteria, therefore, reduces our treatment options for infection and has become a major threat to human health. Resistance was identified soon after the identification of the first antibiotics (2), but throughout the mid-20th century it was kept in check by the development of novel antibiotics. Now, antibiotic resistant infections cause over 20,000 deaths per year in the United States. Even for the over two million antibiotic resistant infections each year that do not lead to death, resistance lengthens treatment time, extending hospital stays, lowering quality of life, and increasing direct healthcare costs by as much as $20 billion (1). Few new antibiotics have been developed since the 1960s (3), meaning that our most successful strategy for fighting antibiotic resistance, finding antibiotics for which resistance has not yet developed, may soon no longer be an option.

Many new strategies are being proposed to combat bacterial infections and slow the spread of resistance, one of which is antibiotic stewardship. Antibiotic stewardship, the practice of using only effective antibiotics and only when necessary, has the potential to slow the spread of resistance regardless of whether other methods are used. To fully practice good antibiotic stewardship will require an improved understanding of the current state of antibiotic resistance in pathogens and predictions of its future state, which will allow those who administer antibiotics to know which ones are likely to be most effective.

Looking to the past spread of antibiotics as a model for what may happen in the future, we have learned that antibiotic resistance has been around since before humankind learned to utilize antibiotics as medicine. Most every antibiotic in use today is, or is derived from, a natural product (3), and bacteria in the environment have been exposed to antibiotics for thousands, if not millions, of years. A study of bacterial DNA found in 30,000 year-old permafrost shows examples of antibiotic resistance genes effective against antibiotics that have been used clinically (4).

This principle extends to bacteria that have evolved in humans. A recent study has shown that bacteria from a group of humans with no prior exposure to medicinal antibiotics encoded antibiotic resistance genes as well (5). It has been proposed that most of these ancient resistance mechanisms originated as self-protection mechanisms in antibiotic producing bacteria (6). Human use of antibiotics on a large scale has selected for resistant bacteria, increasing the prevalence of antibiotic resistance in pathogens and in environmental bacteria. This rise in resistance has been seen by comparing pre-antibiotic strain collections of *Enterobacteriaceae* to modern isolates (7), in increasing rates of methicillin resistant *Staphylococcus aureus* (MRSA) as a percentage of all *S. aureus* (8), and in differences in resistance between antibiotic naïve intestinal bacteria and those that have been exposed to antibiotics (5).

The largest use of antibiotics worldwide is in agriculture, primarily in raising livestock (9). This volume of antibiotic use puts pressure on pathogens and non-pathogens alike to develop resistance, and likely has led to an overall rise in resistance gene prevalence. The use of antibiotics in human medicine is a more acute problem, because it directly pressures human-associated bacteria, including pathogens, to mutate or acquire resistance genes. This has caused hospital-acquired infections (HAIs) to be particularly problematic, as they are highly resistant

and are more likely to affect individuals with a compromised immune system (1). For example, MRSA has largely been associated with HAIs (8), though recently a rise in community acquired MRSA has been seen (10).

To compound these issues, there has been a preference in both medicinal and agricultural antibiotic use for broad-spectrum antibiotics, as they are more likely to be successful even when the exact target bacterium is not known (11). Even appropriate and successful antibiotic treatments, then, create selection pressure for resistance in non-target bacteria. In clinical use, to preserve antibiotics with low prevalence of resistance in pathogens, some antibiotics are characterized as last-resort antibiotics, which reduces selection pressure for resistance to those antibiotics (12). Unfortunately, as resistance rises to the front-line antibiotics, last resort antibiotic must be used more frequently.

Antibiotic resistance can develop vertically, as a result of mutations or insertions, but pathogens can also acquire entirely novel antibiotic resistance genes in a single generation through horizontal gene transfer (HGT), or the exchange of genetic material between bacterial lineages.

There are three major mechanisms by which HGT can occur: transformation, transduction, and conjugation. In transformation, bacteria import free DNA from their environment and incorporate it into their genome. Some bacteria are more likely to import DNA under conditions of cellular stress, such as starvation or DNA damage, which may give them an advantage in adapting to new challenges (13). In transduction, the DNA exchange is mediated by phages, which incorporate an adjacent portion of their host bacterial genome into their phage particles as they enter lytic phase (14). Phage transduction has been shown to carry the cholera

toxin genes found *Vibrio cholerae* strains (15) and the shiga toxin genes found in *Escherichia coli* (16). In conjugation, extra chromosomal plasmids are transferred directly between two bacteria using type-IV secretion systems and dedicated machinery encoded on the plasmids themselves (17). The plasmids involved in conjugation can be larger than 200 kilobases and encode a large number of resistance and other genes (18). Each of the mechanisms of HGT is more likely to occur between closely related organisms, either because the cellular machinery promoting the transfer preferentially targets close relatives, as is the case with transduction and conjugation, or because homologous recombination is required, as is the case with transformation.

Despite the increased likelihood of HGT between close relatives, HGT between bacteria from different phyla has been shown to be enriched in bacteria associated with humans (19) and for anaerobic bacteria in general (20). One way to estimate the HGT of specific genes is using phylogenetic methods, which involve generating phylogenetic trees for the group of bacteria in which HGT is believed to have occurred and for a set of homologous genes from those bacteria. Where the gene and organism trees do not agree, the distance between the gene and organism trees is estimated and a cutoff is applied to differentiate vertically-inherited from horizontally-transferred genes (21). This method applies best to organisms and genes that are fairly closely related, as optimal phylogenetic trees will be easier to identify.

A second way to estimate the HGT of specific genes is though parametric methods, which use characteristics of the DNA sequence of the gene to estimate the likelihood that it evolved in its current genomic context. For example, the codon adaptivity index measures how often a gene uses codons uncommon in the rest of the genome for standard amino acids (22).

Parametric measures are most effective at measuring HGT over large phylogenetic distances, since closely related bacteria are likely to have similar genomic characteristics.

The existence of HGT has several practical effects on the spread of antibiotic resistance in bacteria. First, HGT of resistance genes complicates treatment of bacterial infections because it decouples resistance from phylogeny. Two identical infections caused by very closely related pathogenic strains can have very different antibiotic resistance spectra. Second, HGT allows resistance genes, once evolved, to spread rapidly through bacterial populations, rather than needing to be evolved multiple times. Finally, HGT makes the prevalence of antibiotic resistance genes in environmental non-pathogens problematic, because they can transfer those resistance genes to pathogens they cohabitate with. When the environment is the gastrointestinal tract, pathogens have the opportunity to gain new antibiotic resistance genes from neighboring commensals during the course of infection.

The human gastrointestinal tract is home to a large and diverse bacterial community, hereafter referred to as the intestinal microbiota, which may serve as an antibiotic resistance reservoir for pathogens. Most intestinal microbiota taken from healthy adults are dominated by two phyla in terms of relative abundance: the Firmicutes and the Bacteroidetes (23) and the presence of antibiotic resistance genes in these groups is cause for concern. The Firmicutes, in particular, include a large number of pathogens, such as *Clostridium difficile*, that we currently use antibiotics to treat (24); however, resistance genes in non-pathogenic members of each phyla are still problematic in that they may transfer those resistance genes to a much less abundant group in the intestinal microbiota: the *Enterobacteriaceae*.

In Chapter 3 we explore the role that HGT played in resistance gene acquisition in the human microbiota, particularly exchange between the Firmicutes and the *Enterobacteriaceae*. The *Enterobacteriaceae* includes commensal strains that cause no harm to their human host, but it also includes strains that can cause serious infections, especially HAIs. *E. coli* is perhaps the best known of the *Enterobacteriaceae* and it is a major pathogen. Since 2006, the U.S. Centers for Disease Control and Prevention (CDC) has tracked 21 *E. coli* outbreaks causing intestinal disease (25), but *E. coli* can also cause extraintestinal infections such as urinary tract infections (UTIs) (26), and neonatal meningitis (27). Another *Enterobacteriaceae*, *Klebsiella pneumoniae* frequently causes lung infections in immunocompromised individuals and in patients using respirators, and UTIs (28) as well as organ infections (29). Other common pathogenic *Enterobacteriaceae* include *Enterobacter cloacae*, *Enterobacter aerogenes*, *Klebsiella oxytoca*, *Salmonella* sp., and *Shigella* sp. Many *Enterobacteriaceae* pathogens can survive in a variety of environments, such as in the intestinal microbiota, on surfaces, and on medical devices, which makes them easy to transmit and difficult to eradicate from a hospital once they are endemic.

Virulence alone would not be enough for *Enterobacteriaceae* to maintain their status as major global pathogens given the widespread availability of antibiotics, and many *Enterobacteriaceae* have indeed become highly antibiotic resistant. In its 2013 ranking on antibiotic resistance threats in the United States, the CDC listed *Enterobacteriaceae* twice, depending on specific resistance spectrum (1). The CDC lists Carbapenem resistant *Enterobacteriaceae* (CRE) as one of only three urgent threats (the highest threat level), as these bacteria may be resistant to all antibiotics approved for their treatment. *Enterobacteriaceae* producing an extended spectrum β-lactamase (ESBL) are considered serious threats, the second highest threat level. ESBL-producing *E. coli* and *K. pneumoniae* alone cause over 26,000

infections and 1,700 deaths per year in the U.S. In Chapter 2 we show that pathogenic *E. cloacae* and *E. aerogenes* also can be highly antibiotic resistant, suggesting that the morbitity and mortality rates may be much higher for ESBL-producing *Enterobacteriaceae* as a whole.

What both of the resistance factors that place *Enterobacteriaceae* on the CDC threat ranking have in common is that they are resistant to different subclasses of the same antibiotic class: the β-lactams. The β-lactams are a large and structurally diverse class that have been in use since the beginning of the medical antibiotic era, including the penicillins, the cephalosporins, the cephamycins, the monobactams, and the carbapenems. What all β-lactams have in common is their mechanism of action – irreversible binding of their four-member β-lactam ring to the active site of a bacterial penicillin-binding protein (PBP) – though different β-lactams may preferentially bind different PBPs (30). PBPs are essential enzymes for bacterial growth, with different PBPs catalyzing different steps in cell way synthesis. Because β-lactams are structurally similar to the peptidoglycan substrate of PBPs, they tend to have very wide spectrums of activity, often including gram-positive and gram-negative pathogens. In the *Enterobacteriaceae*, resistance to β-lactams is usually achieved through production of specialized enzymes called β-lactamases that cleave the four-member β-lactam ring, rendering the antibiotic non-functional.

Like β-lactams, β-lactamases are a large and structurally diverse group, and they have been grouped into four structural classes, lettered A-D. Class A includes the TEM, SHV, CTX-M, and KPC sub-classes, and they are the most common ESBLs in pathogenic *Enterobacteriaceae* (31). Unlike the other three classes, class B β-lactamases require binding of metal cations to be active, and are thus also called metallo-β-lactamases. Class C β-lactamases

are expressed from *ampC* genes found in the chromosomes of some *Enterobacteriaceae* species including *E. coli* and *E. cloacae*, though their expression is very limited in that context. Class D genes are also called OXA genes, as they were originally defined by their ability to degrade the β-lactam oxacillin.

Genes from each of these β-lactamase classes can be found both in chromosomes and in plasmids, and each class contains a range of resistance spectra across the β-lactams. Genes providing resistance to the latest generation of β-lactams, the carbapenems, have been found in classes A, B, and D. Often called carbapenemases, these genes are the primary way in which members of the *Enterobacteriaceae* become CRE. A second way is through a combination of an ESBL (usually from either class A or class C) and a deleterious mutation in the outer membrane porin by which the β-lactam enters the periplasm. This second mechanism of carbapenem resistance is not transferrable by HGT, but β-lactamases from all classes are often found on conjugative plasmids. These plasmids mix freely throughout the *Enterobacteriaceae*, and we show in Chapter 2 that recombination between plasmids has created a large diversity of plasmid variants in different strains with differing complements of resistance genes. These plasmids pose a major problem for treatment of infections, not only because they are increasing the prevalence of carbapenems and ESBLs, but also because they link different resistance traits together. Selection pressure for any of the diverse resistance genes on the plasmid will constitute selection pressure for the whole plasmid, increasing the chances of transmission to new organisms.

Researchers and clinicians will need to pursue a suite of approaches to combat the rising tide of antibiotic resistance in organisms such as the Enterobacteriaceae. One necessary component of this armamentarium will be development of new antibiotic classes. One recent

example of this was the discovery of teixobactin, which shows a novel antibiotic mechanism (32). In discovering this compound, researchers also determined that *S. aureus* and *Mycobacterium tuberculosis* did not develop resistance in a laboratory setting over the same time course in which they develop high levels of resistance to other antibiotics (32). Teixobactin does not protect against resistance gained by HGT, though, so if resistance genes exist, or can be evolved in some context, then pathogens may find protection against even this treatment. Because Teixobactin only affects gram-positive bacteria, and because it will take years to optimize and test it for clinical use, other strategies will need to be employed to combat infections with existing antibiotic classes.

Three major strategies have been proposed to keep our current arsenal of antibiotics relevant: synthetic tailoring of antibiotic side groups (33-35), antibiotic combinations (36-38), and antibiotic cycling (39, 40). Each of these strategies is applied to eliminate pathogens, but they often have high collateral damage, disturbing the entire gut microbiota community. Synthetic tailoring is the modification of side groups in an antibiotic molecule to extend its effectiveness or circumvent antibiotic resistance while maintaining the core antibiotic mechanism (34). For example, the β-lactams have undergone several levels of modification since penicillin was first discovered. The effectiveness of penicillin is confined to specific Gram-negative bacteria, but synthetic tailoring generated new antibiotics with expanded activity, such as piperacillin and methicillin (35). The major limitation of synthetic tailoring is that it does not change the fundamental mechanism of the antibiotic. While synthetic tailoring can bypass some types of antibiotic resistance, for others, a single antibiotic resistance gene can give resistance to an entire class of antibiotics. That form of resistance gene is increasing in prevalence. For instance, the recently discovered, plasmid-borne *kpc* and *ndm* genes encode for enzymes that can

degrade all types of β-lactams, and prevalence of these genes continues to increase in hospitals worldwide (41, 42).

Antibiotic combinations have been used to treat organisms for which a single antibiotic treatment is insufficient, and successful combinations often exhibit synergy between the constituents. In a synergistic interaction, the effectiveness of the combination of drugs at a given concentration is greater than the effectiveness of either antibiotic on its own at that concentration. The major advantage of synergistic combinations is that they lower the total drug concentrations needed for killing, which can reduce the toxicity of the treatment to human cells (36, 38). Unfortunately, while lower total concentrations are an advantage in terms of toxicity, they are a disadvantage in terms of evolution of resistance, as bacteria evolve resistance more quickly to synergistic drug combinations than to the drugs used singly (37). This occurs because individual drugs are dosed at sub-therapeutic concentrations, resulting in bacteria facing a lower evolutionary barrier to become resistant to each component of a combination. Once a bacterium evolves resistance to one component, the synergy is broken and other components are no longer at killing concentrations, overall increasing rates of evolution of resistance. This downside has been explicitly demonstrated for two-drug combinations (37), but the theoretical principle may also apply for higher-order synergistic compound combinations.

Antibiotic cycling is an established concept (43-45) and the idea has received renewed recent interest in the context of the phenomenon of collateral antibiotic sensitivity. An antibiotic is considered to confer collateral sensitivity if resistance evolved to that antibiotic makes a bacterium more susceptible to another antibiotic, compared to the wild-type population (40). In some cases, two antibiotics can be reciprocally collateral sensitive, where resistance evolved to either antibiotic increases susceptibility to the other. In this case, it has been proposed that one

antibiotic could be applied until resistance to that antibiotic is manifested and then treatment switched to the other (39). This cycling process could be repeated until the infection is cleared, and since exposure to each antibiotic selects for susceptibility to the other, there would be no net evolution of resistance. This procedure holds much promise, but cycling based on reciprocal collateral sensitivity has yet to be implemented clinically, and it is not known how generalizable collateral sensitivities are between species or even strains. Like Teixobactin, cycling has only been shown to suppress resistance evolved vertically, and could likely be broken by introduction of a resistance genes by HGT.

Separate from specific treatment strategies, increased antibiotic stewardship has remained an important goal for preserving the effectiveness of our current antibiotics. By reducing the selection pressure for antibiotic resistance, it is believed that we can at least slow the spread of resistance to new pathogens. In cases where the resistance gene has a fitness cost in the absence of antibiotics, reduced antibiotic usage may even lead to currently resistant pathogens becoming susceptible again. Since resistance will eventually evolve against any antibiotic therapy, it will be important to implement stewardship regardless of the other treatment strategies taken.

Unfortunately, though stewardship is simple in theory, there are practical challenges. The first is that much of the widespread use of antibiotics in agriculture is not used for treating infections but rather prophylactically (46) or to promote growth (47). These applications provide tangible benefits, but they also create a continual selection pressure for resistance in bacteria with close ties to the human population. Even in directly treating human disease up to 50% of antibiotic treatments are unnecessary or improper (1), and reducing that number will require technological advances.

The crux of the problem is that it takes approximately two days from the time an isolate is brought to a clinical laboratory to identification of its resistance spectrum, using current methods (48). There is an inverse relationship between time taken to administer the proper treatment and patient outcomes (49), so clinicians will prescribe antibiotics based on empirical observations of similar infections in that geographic area. Because of HGT, infections that present with identical characteristics and are caused by closely related pathogens may have very different resistance profiles.

This leads to two possibilities for sub-optimal antibiotic choice: 1) the pathogen is resistant to the empirically chosen antibiotic and the treatment fails, or 2) the antibiotic was effective, but a more front-line antibiotic would have been equally effective. Both cases promote increased resistance. The first case will require follow-up with a second antibiotic course, increasing the total antibiotics used, and the second case will increase selection pressure for last-resort antibiotics. As the prevalence of antibiotic resistance increases, both types of errors will become more common, leading to a cycle of increasing resistance in human-associated bacteria, including pathogens. To slow this cycle of resistance for the vital β-lactamase antibiotic class in the *Enterobacteriaceae*, I used genomic analysis to identify the patterns of resistance gene transfer with the *Enterobacteriaceae*, and I developed and evaluated software that can reduce the time taken to provide pathogen resistance profiles to clinicians.

My first objective was to determine if increases in β-lactamase prevalence within the *Enterobacteriaceae* is primarily driven by vertical inheritance or by HGT. In Chapter 2, we look at this question through the lens of two carbapenemase genes, *kpc* and *ndm-1*. These genes had each recently been discovered in pathogenic *Enterobacteriaceae*, and the areas where they were

endemic were geographically separated. By looking at high-resolution phylogenies of the isolates that encoded each gene we could determine if they were from globally disseminated strains, or local strains that were undergoing clonal expansion as a result of their resistance. We also sequenced the plasmid DNA from isolates encoding either carbapenemase to determine if HGT of these genes was being driven by one or two consistent plasmids, or if they were a population of plasmids with some similar elements. I performed all of the sequence analysis for this publication, as well as generating the figures and being the primary author of the text.

I was next interested in determining the potential for HGT of resistance genes from other members of the microbiota to the *Enterobacteriaceae*. In Chapter 3 we use the Resfams antibiotic resistance gene database (50) to identify all of the antibiotic resistance genes present in a mock intestinal microbiota. We use a parametric tests for HGT on all of the resistance genes from this community and from a set of previously sequenced intestinal bacteria to estimate the effects that HGT over long phylogenetic distances has had on the prevalence of antibiotic resistance in the microbiota. We then use functional metagenomics to identify the genes from this same mock community that could be utilized by *E. coli* for antibiotic resistance. My co-first author and I equally divided the experimental, analysis, and writing work.

Having explored the effects that HGT can have on resistance gene composition in the microbiota, I turned to a method for improving antibiotic stewardship and slowing the spread of resistance: faster diagnostics for antibiotic resistant infections. It has been suggested that whole genome sequencing could be used to provide faster resistance diagnostics, particularly because of its utility in other clinical applications (48). In Chapter 4 we evaluate the effectiveness of two algorithms in correctly translating genome sequence into susceptibility profiles for clinical isolates. We also identify and quantify the sources of error in each algorithm, so that further

improvements can be made. I performed all of the analysis in this paper, including implementing both algorithms, and was the primary author of the text and figures.

Chapter 2: KPC and NDM-1 are harbored by related Enterobacteriaceae strains and plasmid backbones from distinct geographies

Mitchell W. Pesesky*[1], Tahir Hussain*[1,2], Meghan Wallace[3], Bin Wang[1,3], Saadia Andleeb[2], Carey-Ann D. Burnham[3,4], Gautam Dantas[1,3,5‡].

[1]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, United States of America.

[2]Atta ur Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad, Pakistan

[3]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, United States of America.

[4]Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri, United States of America

[5]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri, United States of America.

*These first authors contributed equally to this work.

2.1 Abstract

To characterize the genomic context of New Delhi Metallo-β-lactamase-1 (NDM-1) and *Klebsiella pneumoniae* Carbapenemase (KPC), we sequenced 78 *Enterobacteriaceae* isolates encoding KPC, NDM-1, or no carbapenemase from Pakistan and the United States. We found examples of NDM-1 and KPC expressing strains and plasmids with high similarity to each other.

2.2 Background

Pathogenic *Enterobacteriaceae*, including *Escherichia coli* and *Klebsiella pneumoniae*, are a major cause of multi-drug resistant (MDR) infections in hospitals worldwide. They have

recently acquired resistance to the carbapenems, and the CDC named carbapenem resistant *Enterobacteriaceae* as one of the three most urgent MDR threats (1). In the *Enterobacteriaceae*, β-lactam resistance, including carbapenem resistance, is primarily caused by enzymatic degradation by β-lactamases. Two carbapenemase subclasses are especially problematic: *Klebsiella pneumoniae* carbapenemase (KPC) and New Delhi Metallo-β-Lacatamase-1 (NDM-1). KPC, identified in 2001 (51), has become endemic in several non-contiguous areas of the world, including the U.S., Israel, Greece, South America, and China (52). NDM-1 was first described in 2008, though retrospective studies identified NDM-1 from 2006 (53), and to be abundant in New Delhi water samples (54). Most patients from whom NDM-1 is isolated have an epidemiological link to the Indian subcontinent, but it has also recently become endemic in the Balkans and Middle East (55).

The spread of antibiotic resistance (AR) genes such as NDM-1 and KPC is facilitated by horizontal gene transfer (HGT) between bacteria (56). Using HGT, globally disseminated pathogens combine the most effective AR genes from diverse geographies into multi-drug resistance plasmids, which themselves spread between strains. Recombination and transposition have created populations of these plasmids with related architectures, but varying in their composition of antibiotic resistance cassettes (31). This has enabled both KPC and NDM-1 to rapidly expand within the *Enterobacteriaceae* and other Proteobacterial pathogens such as *Acinetobacter baumanii* (57, 58). AR genes can also spread through clonal expansion in successful pathogenic strains, for example KPC in *K. pneumoniae* sequence type (ST) 258 (18), and the extended-spectrum β-lactamase CTX-M-15 in *E. coli* ST131 (59). Both HGT and clonal expansion have allowed KPC and NDM-1 to rapidly spread to distant locations after their emergence (31, 55). The similarities in the spread and resistance spectra of KPC and NDM-1

(both provide resistance to nearly all β-lactam antibiotics), leads to the hypotheses that similar mobile elements will make both genes available to similar pathogen populations. We tested this hypothesis by examining clinical *Enterobacteriaceae* isolates encoding NDM-1, KPC, or no carbapenemase from Pakistan and the United States, respectively.

2.3 Materials and Methods

Summary: We collected 450 Pakistani bacterial isolates (including 195 *Enterobacteriaceae*) (PH) between February 2012 and March 2013 from Pakistan Railway General Hospital, Rawalpindi and the Pakistan Institute of Medical Sciences in Islamabad. From this collection we randomly selected 55 *Enterobacteriaceae* for whole genome sequencing. We then selected 23 US isolates (WU) between January 2010 and June 2013 from Barnes Jewish Hospital in Saint Louis, Missouri to have similar proportions of β-lactam susceptibility and resistance to the PH isolates for sequencing. All isolates were de-identified and retrieved from existing strain banks. The combined set includes 33 *E. coli*, 30 *K. pneumoniae*, 9 *Enterobacter cloacae*, and 6 *Enterobacter aerogenes* (Table 2.1). We extracted plasmid DNA from 9 isolates encoding NDM-1, 11 isolates encoding KPC, and 3 isolates encoding CTX-M-15, and performed shotgun sequencing on those plasmid preparations.

Sample Selection, Processing, and Phenotyping. We collected 450 Pakistani bacterial isolates (PH) initially recovered from de-identified clinical samples from urinary, blood stream, genitourinary, and wound infections collected between February 2012 and March 2013 at Pakistan Railway General Hospital, Rawalpindi, Pakistan, and the Pakistan Institute of Medical Sciences in Islamabad, Pakistan. These included all ESKAPE pathogen isolates available in the Pakistani hospital strain banks during the indicated collection period. From these 450 isolates, we chose a random subset of 55 isolates from the *Enterobacteriaceae* family (from a total of 195

Enterobacteriaceae in this collection) for phenotypic and genotypic analysis. We also selected 48 US *Enterobacteriaceae* isolates (WU) from banked, de-identified frozen stocks of *Escherichia coli*, *Klebsiella pneumoniae*, and *Enterobacter cloacae* at Barnes Jewish Hospital/Washington University School of Medicine in Saint Louis, Missouri, United States to have beta-lactam resistance and susceptibility phenotypes in similar proportions to the Pakistani *Enterobacteriaceae* isolates, with a particular focus on the meropenem resistance phenotype (protocols for growth and phenotyping are described below). The WU strains were originally isolated from urine, respiratory, bone, and bile specimens between January 2010 and June 2013. Of the 48 WU isolates, 23 isolates were chosen for genome sequencing to generally match the species distribution of the 55 PH isolates as well as their beta-lactam resistance profiles, with the exception of *Enterobacter aerogenes*, for which none were available in the WU collection. In total, 33 *Escherichia coli* (24 PH, 9 WU), 30 *Klebsiella pneumoniae* (19 PH, 11 WU), 9 *Enterobacter cloacae* (6 PH, 3 WU), and 6 *Enterobacter aerogenes* (all 6 PH) isolates were included for the whole genome sequencing analysis (Tables 2.1 and 2.2). We cultivated all isolates on MacConkey and sheep's blood agar (Hardy Diagnostics). We then grew single colonies in LB broth liquid culture for DNA extraction. We assessed each isolate for susceptibility to ampicillin, cefazolin, cefotetan, ceftazidime, ceftriaxone, cefepime, meropenem, ciprofloxacin, trimethoprim-sulfamethoxazole, gentamicin, doxycycline, and chloramphenicol by Kirby-Bauer disk diffusion according to Clinical and Laboratory Standards Institute guidelines and interpretive criteria (12). Prior to whole genome sequencing, the species identity of PH and WU isolates was determined with VITEK MS MALDI-TOF MS v2.0 knowledgebase (bioMerieux) as previously described (60, 61). We then extracted total DNA using the Invitrogen Charge Switch gDNA Mini Bacteria kit per the manufacturer's protocol. We also extracted

plasmid DNA from 11 KPC and 9 NDM-1 encoding isolates (as well as 3 CTX-M-15 encoding isolates), as determined by PCR and the genome sequencing, using the Qiagen Large Construct kit per the manufacturer's protocol. We included one non- *Enterobacteriaceae* plasmid preparation from an *Acinetobacter baumannii* isolate (PH), which had been identified to contain NDM-1 by PCR.

Illumina Library Preparation. We sheared 500ng of total DNA from each isolate to ~300 bp fragments fragments in nine rounds of shearing of ten minutes each on the BioRupter XL. In each round the power setting was 'H' and samples were treated for 30s and allowed to rest for 30s. Each sample was concentrated using the Qiagen MinElute PCR purification kit per the manufacturer's protocol. End Repair of the sheared DNA fragments was initiated with the addition of 2.5 μl of T4 DNA ligase buffer with 10mM ATP (NEB, B0202S), 1 μl of 1 mM dNTPs (NEB), 0.5 μl T4 Polymerase (NEB, M0203S), 0.5 μl T4 PNK (NEB M0201S), and 0.5 μl Taq Polymerase (NEB, M0267S). This mixture was incubated at 25°C for 30 min, then at 75°C for 20 min. Barcoded adapters were then added to the solution along with 0.8μl of T4 DNA ligase (NEB, M0202M), for the purpose of ligating the adapters to the DNA fragments. This solution was then incubated at 16ºC for 40min, then 65°C for 10min. At this point the adapter-ligated DNA was purified using the Qiagen MinElute PCR purification kit per the manufacturer's protocol.

The DNA fragments were then size selected on a 2% agarose gel in 1X TBE buffer stained with Biotium GelGreen dye (Biotium). DNA fragments were combined with 2.5uL 6X Orange loading dye before loading on to the gel. Adaptor-ligated DNA was extracted from gel slices corresponding to DNA of 250-300bp using a QIAGEN MinElute Gel Extraction kit per the

manufacturer's protocol. The purified DNA was enriched by PCR using 12.5µL 2X Phusion HF Master Mix and 1µL of 10µM Illumina PCR Primer Mix in a 25µL reaction using 1µL of purified DNA as template. DNA was amplified at 98°C for 30 seconds followed by 18 cycles of 98°C for 10 seconds, 65°C for 30 seconds, 72°C for 30 seconds with a final extension of 5min. at 72°C. Afterwards, the DNA concentration was measured using the Qubit fluorometer and 10nmol of each sample (up 106 per lane of sequencing) were pooled. Subsequently, samples were submitted for Illumina HiSeq-2500 Pair-End (PE) 101bp sequencing at GTAC (Genome Technology Access Center, Washington University in St. Louis) at 9pmol per lane.

Genome sequence assembly. All sequencing reads were de-multiplexed by barcode into separate genome bins. Reads were quality trimmed to remove adapter sequence and bases on either end with a quality score below 19. Any reads shorter than 31bp after quality trimming were not used in further analysis. The best reference sequence was chosen for each isolate or plasmid by mapping 10000 reads chosen randomly from that isolate against all reference genomes (from NCBI Genome, downloaded July 14th 2014) of the same species as the isolate (in the case of genomic DNA assembly) or against all plasmid sequences containing NDM-1, KPC, or CTX-M (in the case of plasmid DNA assembly). Reads were mapped using Bowtie 2(62) (command: bowtie2 –x <reference_genome_index_name> -1 <forward_read_file> -2 <reverse_read_file> -q --phred33 --very-fast –I 100 –X 600 --no-discordant --no-mixed --no-unal --no-hd --no-sq --omit-sec-strand). The genome or plasmid against which the highest percentage of reads mapped was used as the reference sequence for that assembly. It was empirically determined that if this first mapping included fewer than 60% of the reads, then the assembly would be best done completely *de novo*. For isolates with >60% of reads matching a reference sequence, all reads were mapped to that sequence (command: bowtie2 –x

<reference_genome_index_name> -1 <forward_read_file> -2 <reverse_read_file> -q --phred33 --very-sensitive-local -I 200 -X 1000 -S <sam_output>). Variants from the reference were called using samtools (commands: samtools view -buS <sam_file> | samtools sort -m 4000000000 - <sample_prefix> ### samtools index <bam_file> ### samtools mpileup -uD -f <reference_genome> <bam_file> | bcftools view -bcv - > <bcf_file> ### bcftools view <bcf_file>). The variant call format file was then filtered to remove SNPs with a quality score lower than 70 or coverage greater than twice the average coverage expected per base. Custom scripts were then used to extract DNA sequences from the reference genome with > three independent reads, to create a fragment file of regions in the sample genome matching the reference genome modified with high-quality variant information.

De novo assembly of the reads from each isolate was completed using Velvet(63) (commands: velveth <output_directory> 51 -fastq -shortPaired <interleaved_reads> ### velvetg <output_directory> -ins_length 400 -exp_cov <kmer_coverage> -cov_cutoff <coverage_cutoff>). Kmer coverage was calculated as: total read coverage*0.50 (because the kmer length was approximately half the read length), and the coverage cutoff was calculated as the kmer coverage divided by eight. If a complete reference mapping was performed, then contigs from the *de novo* assembly and reference mapping were put in an additional velvet assembly step as long reads with the original reads files (commands: velveth <output_directory> 51 -fastq -shortPaired -separate <forward_reads> <reverse_reads> -fasta -long <de_novo_fragments> <reference_fragments> ### velvetg <output_directory> -ins_length 400 -clean yes -conserveLong yes -exp_cov <kmer_coverage> -cov_cutoff <coverage_cutoff> -scaffolding yes -long_mult_cutoff 0). Finally all fragments were collapsed on nucleotide identity using cd-hit (command: cd-hit-est -I <fragment_file> -o <collapsed_file> -d 0 -M 0). All

fragments smaller than 500bp were partitioned to a separate file by a custom script. Plasmid sequences were assembled by this same method, with the sequences of all complete plasmids encoding, NDM-1, KPC, or CTX-M used as references.

ORF prediction and annotation. ORF prediction for each genome was performed separately using GeneMark(64) models based on the closest reference genome (command: gmhmmp -m <model_name> -o <outfile> -a <contig_name_file>). Each ORF was compared to three databases of profile hidden Markov models using HMMR(65): Pfam (command: hmmscan --cut_ga -o /dev/null --tblast <target_out_file> --domtblast <domain_out_file> <Pfam_database_file> <protein_input_file>), TIGRFAMs (command: hmmscan --cut_ga -o /dev/null --tblast <target_out_file> --domtblast <domain_out_file> <database_file> <protein_input_file>), and Resfams (dantaslab.wustl.edu/resfams/) (50) (command: hmmscan --cut_ga -o /dev/null --tblast <target_out_file> --domtblast <domain_out_file> <database_file> <protein_input_file>). All functional annotations were concatenated into a single file by a custom script. 756 *E. coli* and 54 *K. pneumoniae* completed and draft genomes were downloaded from the National Center for Biotechnology Information (NCBI) on April 15th 2014, for the purpose of comparing to the isolate set. GeneMark models from the completed genomes were used to predict ORFs for those genomes, while for draft genomes models created from *E. coli* K12 MG1655 (for *E. coli* draft genomes) or *K. pneumoniae* KCTC 2242 (for *K. pneumoniae* draft genomes) were used for ORF prediction. All genome and plasmid sequences were deposited into NCBI (BioProject accession number: PRJNA261540).

In silico MLST. Multi-Locus Sequence Typing **(**MLST) profiles were downloaded from PubMLST (pubmlst.org). When an absolute MLST profile could not be identified for an organism (because of ambiguous bases or incomplete assembly of one or more loci) the

remaining possible sequence types (ST) based on the incomplete information were identified. In all cases a strain could be identified as one of at most 19 ST using this methodology. MLST profiles were only applied to *Escherichia coli* and *Klebsiella pneumoniae* since the PubMLST database does not contain an *Enterobacter* table.

Core genome alignment. Command for whole genome alignment using mugsy(66): command: mugsy --directory <output_directory> --prefix <output_prefix> <genome_fasta_1> <genome_fasta_2> … <genome_fasta_N>. For *E. coli*, *K. pneumoniae*, and *E. cloacae* a single reference genome was included in the alignment to provide context (*E. coli* K12 MG1655, *K. pneumonaie* KCTC 2242, and *E. cloacae* ATCC 13047 respectively). Poorly aligned regions (i.e. plasmids or recombined regions, which could create noise in the phylogenetic signal) were removed using Gblocks(67) (command: Gblocks <input_file> -t=d -b3=24) leaving only the core genome alignment. Maximum likelihood trees made made by RaxML(68) (command: raxmlHPC-SSE3 -s <input_file> -n <output_file> -m GTRGAMMA -d -f a -N 100 -x 54321 -w <output_directory>) and FastTree(69) (command: FastTree -gtr -nt -gamma -nome <input_file> > <output_file>) were compared for agreement. When trees made by both methods were in agreement, the output from FastTree was used for visualization. Files were converted between various required formats by custom scripts.

Subspecies clades were defined as groups of branches descended from a common ancestor where no individual branch within the clade could have more than 0.005 substitutions per site. This definition yielded the same groupings as the *in silico* MLST described above, in all cases where all members of a clade could be assigned to a known ST (Figure 2.S1A-B).

Specific β-lactamase identification. A BLAST database was constructed from the amino acid sequences of all β-lactamases cataloged in the Bush and Jacoby(70) database at www.lahey.org/Studies/ (accessed March 25th, 2014). Genes from our genomes annotated as β-lactamases were extracted and compared against this database by BLAST. Exact matches were then re-annotated with their specific β-lactamase name, while inexact matches were recorded as their closest hit plus an asterisk.

Plasmid Comparisons. We compared plasmid sequences by an all-against-all pairwise nucleotide BLAST alignment. For each pair of plasmids, we calculated the percentage of each plasmid that aligned at >99% identity. We then binned the percentages from each pairwise alignment into groups based on the defining β-lactamase of their query and subject plasmids. We also generated network diagrams from the pairwise BLASTs using custom Python scripts and Cytoscape (71), only including regions above 99% identity and over 500 bp.

2.4 Results

The sampled Enterobacteriaceae isolates are phylogenetically diverse and include multiple examples of known pathogenic sequence types. We performed WGS of each isolate, totaling 33 *Escherichia coli* isolates, 30 *Klebsiella pneumoniae* isolates, 9 *Enterobacter cloacae* complex isolates, and 6 *Enterobacter aerogenes* isolates. We then used whole genome alignment of the core genomes of each species to reconstruct the phylogenetic relationships of each isolate at high resolution (Figure 2.S1). The species trees demonstrate that we sampled genomes from a variety of evolutionary clades as well as from multiple members of specific clades. They also demonstrate that clades could include isolates from both the United States and Pakistan, allowing us to ignore the geographic variable and group the isolates by carbapenemase carriage for subsequent analyses. We also used housekeeping gene sequence from each isolate to perform in

silico MLST, allowing us to compare our phylogenetic analysis to previously identified sequence types. We found that the clades on our tree include globally-disseminated pathogen sequence types, such as ST131 in *E. coli* and ST11 (single locus variant of ST258) in *K. pneumoniae*.

ST131 is noted for its virulence as well as for its frequent association with the CTX-M β-lactamases and fluoroquinolone resistance (59, 72, 73). Previous reports have found ST258, and closely related *K. pneumoniae*, to have relatively high rates of carbapenemase carriage (18, 74). Despite the utility of MLST-based classification for large-scale epidemiological purposes, binning clinical isolates into sequence types masks genotypic and phenotypic variation due to HGT or single nucleotide polymorphisms, and therefore MLST cannot be used for fine-grained epidemiology or as an accurate predictor of antibiotic susceptibility. For example, two previous studies have shown that ST131 can be subdivided into three distinct lineages with different rates of antibiotic resistance (59, 72). One of these studies found that the rapid global expansion of ST131 has been driven by the success of a specific subclone of ST131 that encodes fluoroquinolone resistant *gyrA* and *parC* alleles and CTX-M-15 (72), a characterization which fits 7 of our 11 ST131 isolates. We also identified a single ST131 isolate carrying KPC-2, which was resistant to all β-lactams tested. We also observed *K. pneumoniae* ST11 isolates carrying KPC-3, and others carrying NDM-1, which fits with reports characterizing ST11 as being highly common worldwide and frequently encoding carbapenemases (18, 74).

Antibiotic resistance phenotypes. To establish the overall susceptibility profiles of each of our strains, we performed phenotypic tests using Kirby Bauer Disk diffusion in accordance with CLSI guidelines on all 78 clinical isolates against 12 antibiotics including 7 β-lactams (Table 2.1). We found that 63% of all isolates were resistant to ciprofloxacin, a fluoroquinolone commonly used to treat urinary tract infections. We also found resistance to trimethoprim-

sulfamethoxazole in 65% of isolates, and gentamicin, doxycycline, and chloramphenicol exhibited *in vitro* resistance in 45%, 54%, and 56% of isolates, respectively. In the β-lactams, we saw near universal resistance to ampicillin (96% of isolates) and variable resistance to the cephalosporins. A high rate of resistance to meropenem was observed (31% of isolates), but this finding was not surprising since this was the property on which many of the isolates had been selected.

Using AR gene predictions from the Resfams database (75) and core genome alignment, we constructed a phylogenetic tree for each species in our set, overlaid by the β-lactamases encoded by each isolate (Figure 2.1). Isolates from both locations were found to be members of the same subspecies clades (Figure 2.S1) and to contain similar repertoires of β-lactamases (Figure 2.1), indicating geography is not a discriminating variable for these isolates. Many of these isolates were also MDR, with resistance to ciprofloxacin, trimethoprim-sulfamethoxazole, gentamicin, doxycycline, and chloramphenicol occurring in 63%, 65%, 45%, 54%, and 56% of isolates respectively. As expected from previous work, *E. coli* ST131 isolates had high rates of CTX-M carriage (82%, Figure 2.1A) and ciprofloxacin resistance (100%).

The variety of strains that we discovered encoding KPC and NDM-1 is consistent with existing evidence that HGT is a major factor in their spread. All KPC genes were proximal to Tn4401 and all NDM-1 genes were carried on ISAba125, mobile elements with which each gene has respectively been previously associated (76). We observed multiple examples of NDM-1 within the *K. pneumoniae* ST11 clade (74) (Figure 2.1B, Figure 2.S1B), a close relative of ST258. This could be caused by clonal expansion or multiple HGT events, and highlights that lineages known to encode KPC are now acquiring NDM-1 as well. We also observed high rates of NDM-1 carriage in *Enterobacter* isolates (Figure 2.1C-D), which in general showed a high

number (maximum: 8) and wide variety of β-lactamases. These isolates were also MDR, with 57% of the *Enterobacter* isolates resistant to all or all but one of the antibiotics tested. At best these *Enterobacter* strains are a reservoir for resistance in Pakistan, at worst they are the vanguard of an expansion of carbapenem-resistant *Enterobacter* infections.

Previous observations have predominantly found KPC and NDM-1 to be expressed from plasmids (18, 55). To characterize the sequence similarity of plasmids within the NDM- and KPC-carrying plasmid populations, we purified and sequenced plasmid DNA from 9 isolates encoding NDM-1, 11 encoding KPC, and 3 encoding CTX-M-15. Sequencing revealed that these plasmids include representatives from IncHI2, IncY, IncN, IncFIA, IncFIB, IncFIC, and IncI1 incompatibility groups. Using reciprocal BLAST alignment between each pair of plasmid preparations, we calculated the percentage of each plasmid shared using a 99% identity threshold. We performed this same analysis for all sequenced plasmids containing NDM-1, KPC, or CTX-M available in the National Center for Biotechnology Information (NCBI, accessed February 10th, 2014) together with our set (Figure 2.2) and separately (Figure 2.S2). Certain components, primarily mobile elements, were abundant within these plasmids: the average plasmid shared 500 contiguous bases with 58 of the other plasmids; however, median BLAST identity for this pairwise comparison was below 12%, even when considering plasmids with the same β-lactamase, suggesting that both carbapenemases exist within a variety of plasmid configurations.

To visualize this comparison of carbapenemase plasmids, we generated a network diagram where each node (circle, triangle, or chevron) represents a plasmid, and each line represents shared sequence between two plasmids (Figure 2.2B). Node size and line width correlate to the number of nucleotides contained in the plasmid or sharing interaction. This

visualization shows the abundant small, shared regions that exist between most plasmid pairs, represented as thin background lines. This visualization also highlights the larger shared regions that indicate highly similar plasmids, represented by the few wide lines. These outliers were often between pairs of plasmids encoding the same β-lactamase, but were also observed between NDM-1 and KPC containing plasmids (maximum: 79% of smaller plasmid length).

2.5 Conclusion

Together, this evidence supports our hypothesis that the same strains that acquire one carbapenemase will also have access to the other. We also observed a few of instances of NDM-1 and KPC existing in highly similar plasmids from both our set and from finished plasmids in NCBI, though plasmid diversity was under-sampled in both. Given the similarity of carbapenemase-negative strains to those carrying KPC or NDM-1 and the high diversity of plasmids in which KPC and NDM-1 can be found, we anticipate that global carbapenem usage will drive HGT of both of these carbapenemases into additional strain and plasmid backgrounds.

2.6 Figures

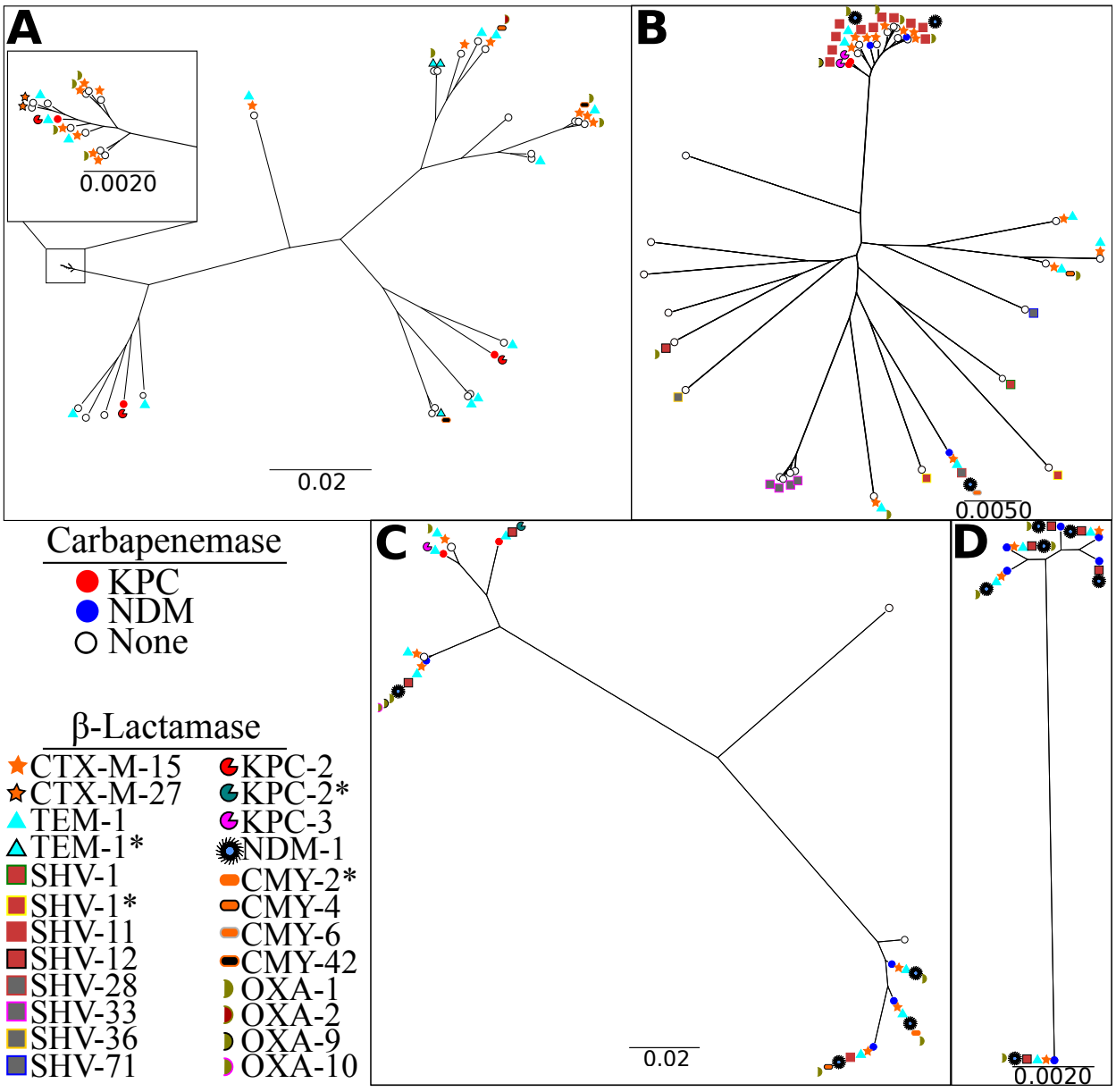
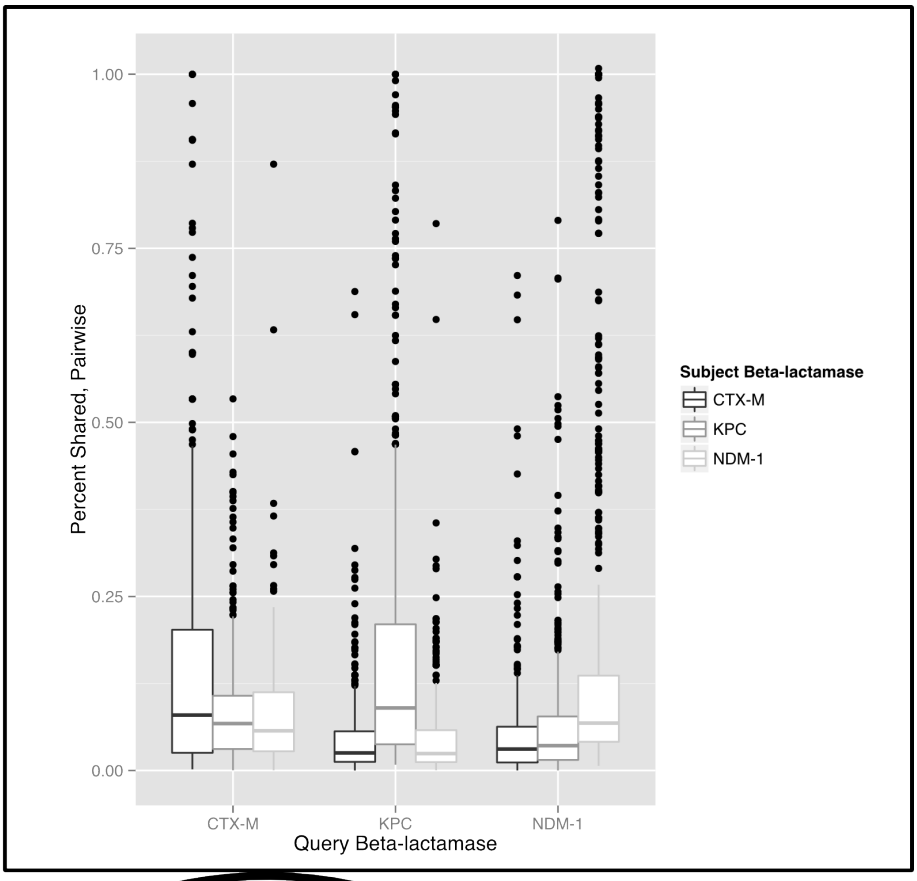


Figure 2.1

Distribution of antibiotic resistance genotypes across isolate phylogenies. Phylogenetic trees from Figure 2.S1 have been annotated with the specific β-lactamases encoded by those isolates in lines extending from the isolate leaf. Circles at each leaf are colored to represent carbapenemase carriage. A) *Escherichia coli,* B) *Klebsiella pneumoniae,* C) *Enterobacter*

cloacae, D) *Enterobacter aerogenes*. *Denotes an unnamed single nucleotide variant of the named β-lactamase.


1.00
0.75
0.50
0.25
0.00
Percent Shared, Pairwise
CTX-M
KPC
NDM-1
Query Beta-lactamase
Subject Beta-lactamase
CTX-M
KPC
NDM-1

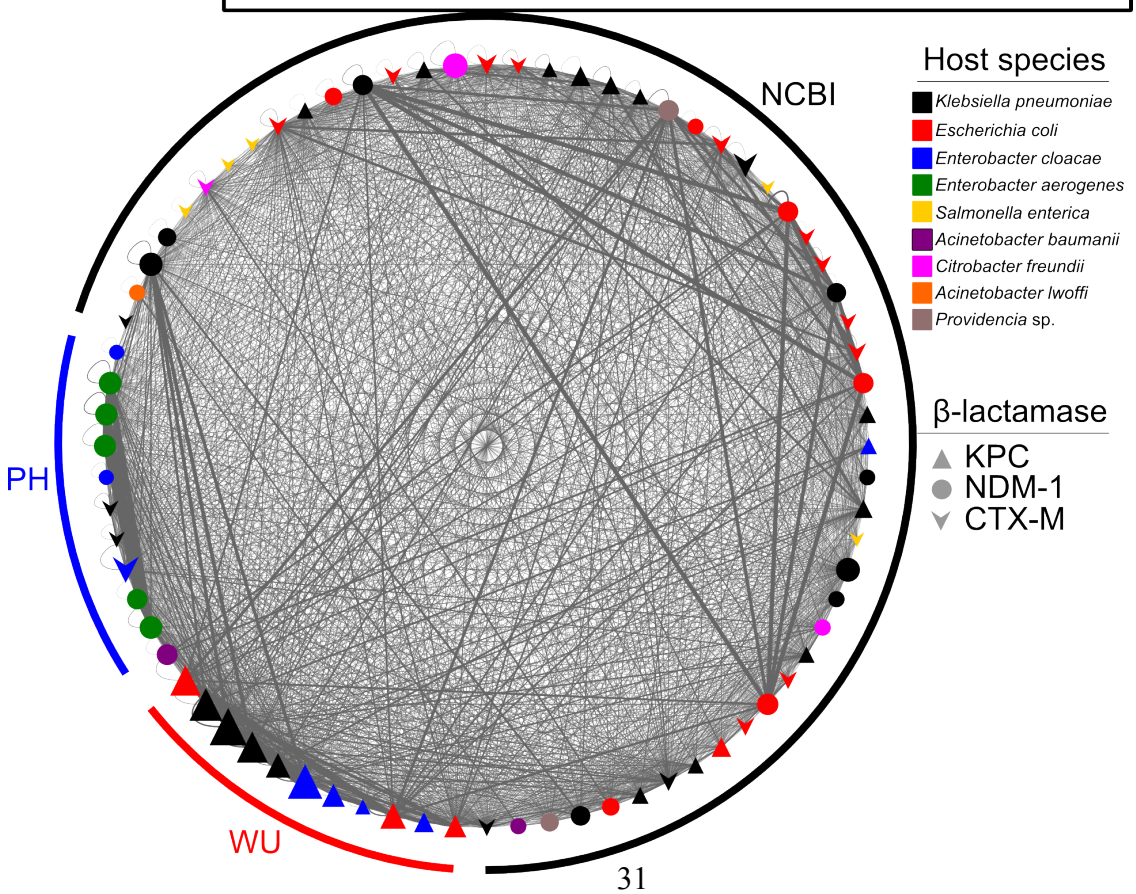

NCBI
PH
WU
Host species
Klebsiella pneumoniae
Escherichia coli
Enterobacter cloacae
Enterobacter aerogenes
Salmonella enterica
Acinetobacter baumanii
Citrobacter freundii
Acinetobacter lwoffi
Providencia sp.
β-lactamase
KPC
NDM-1
CTX-M

Figure 2.2

Pairwise BLAST identity of all CTX-M, KPC, and NDM-1 plasmids from PH and WU plasmid preparations, and NCBI complete plasmids. An all-against-all plasmid BLAST was performed and plasmid interactions were defined by the percentage of the query plasmid conserved (at ≥ 99% identity) in the subject plasmid. A) plasmid interactions were collected based on the defining β-lactamase of their query and subject plasmids. Box and whisker plots represent the range of pairwise sharing values within this population of plasmids, with the upper and lower boundaries of the box corresponding to the 1st and 3rd quartiles, the "whiskers", or error bars, representing 1.5 times the interquartile range, and points representing outliers beyond the whiskers. B) A network map in which nodes (triangle, circle, or chevron) represent individual plasmids and lines represent regions shared between plasmids, with line width proportional to the number of nucleotides contained in fragments > 500bp in length at >99% sequence identity. Genetic elements repeated within the same plasmid DNA are represented by lines that leave and return to the same node. Plasmid sequence origin is indicated in arcs around the network.

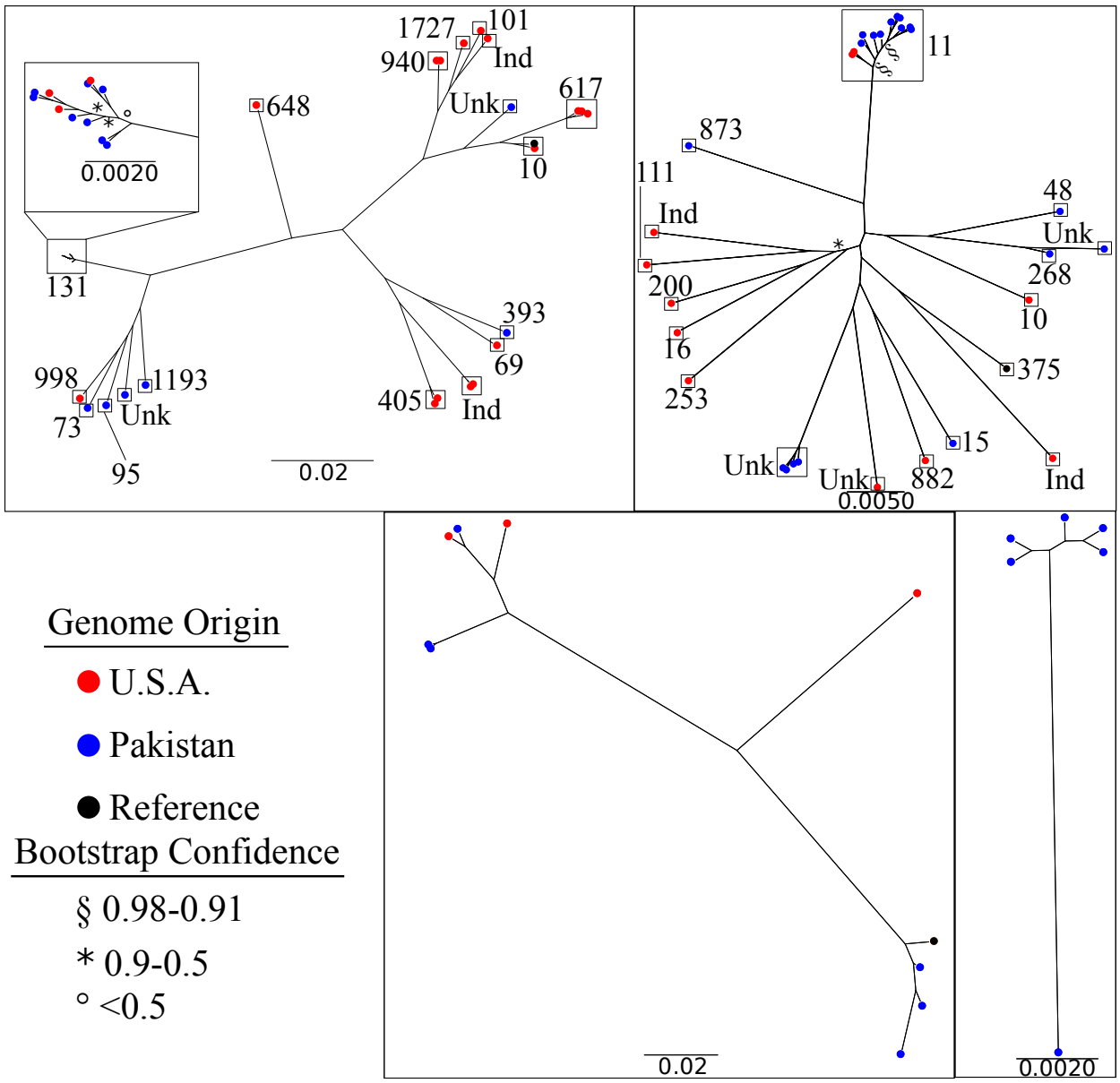


Figure 2.S1

Phylogenetic trees for PH and WU samples. Trees are separated by species, A) *Escherichia coli*, B) *Klebsiella pneumoniae*, C) *Enterobacter cloacae*, and D) *Enterobacter aerogenes*, but not rooted. Bootstrap values are 1 for each branch unless otherwise noted. For each species approximately 50% of the genome was determined to be core, and was used for phylogenetic inference. Scale bars indicate the nucleotide substitutions per site. In a) and b) sequence types

(ST) as determined by *in silico* multi-locus sequence typing are indicated by boxes grouping members of the same ST together. Unk = ST does not correspond to any reported in pubMLST, Ind = exact ST could not be determined due to sequencing error. Reference genomes included for *E. coli* (K12 MG1655), *K. pneumoniae* (KCTC 2242), and *E. cloacae* (ATCC 13047) on their respective trees.

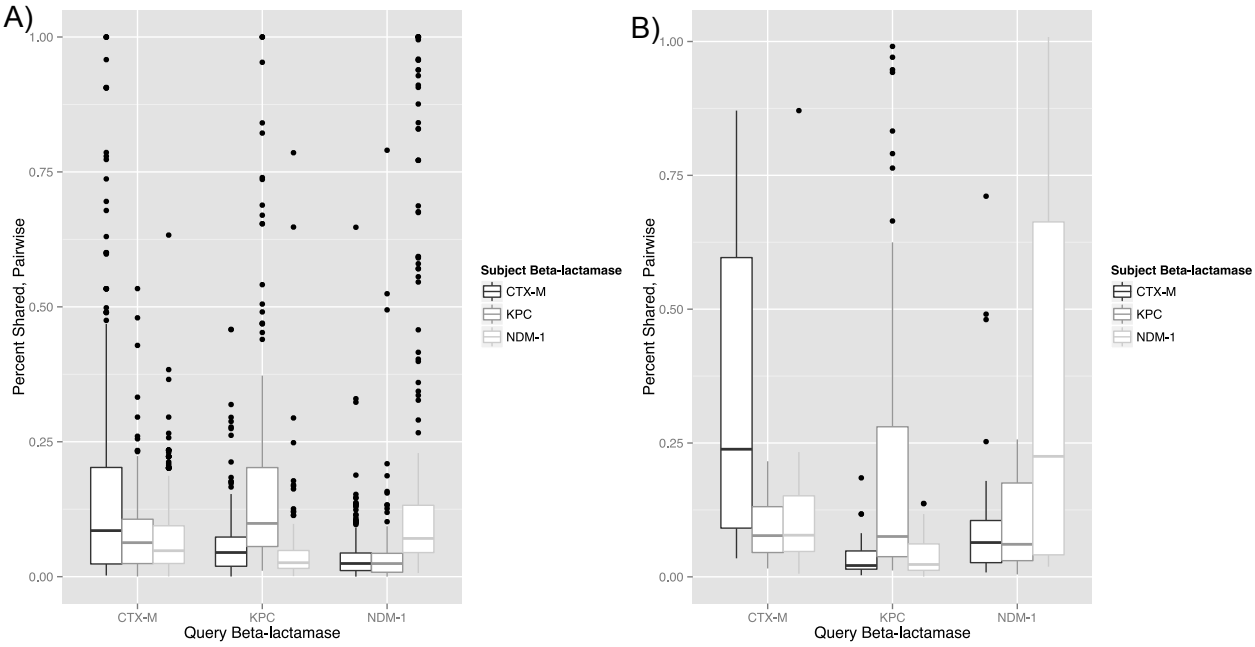


Figure 2.S2

Sequence conservation between plasmids containing NDM-1, KPC, or CTX-M β-lactamases. All plasmids from A) NCBI and B) this study that contained a NDM-1, KPC, or CTX-M β-lactamase were analyzed by all-against-all BLAST. Plasmid interactions were defined by the percentage of the query plasmid conserved (at > 99% identity) in the subject plasmid. Plasmid interactions were plotted based on the defining β-lactamase of their query and subject plasmids.

2.7 Tables

Species	Count.	Pheno-type	AM	CZ	CTT	CAZ	CRO	FEP	MEM
E. coli n=9	U.S.A.	Res	78%	67%	22%	44%	44%	44%	44%
		Int	0%	11%	22%	0%	0%	0%	0%
		Sus	22%	22%	56%	56%	56%	56%	56%
E. coli n=24	PK	Res	96%	83%	13%	38%	63%	21%	0%
		Int	4%	17%	0%	13%	0%	21%	0%
		Sus	0%	0%	88%	50%	38%	58%	100%
K. pneumoniae n=11	U.S.A.	Res	100%	36%	0%	36%	36%	36%	36%
		Int	0%	0%	36%	0%	0%	0%	0%
		Sus	0%	64%	64%	64%	64%	64%	64%
K. pneumoniae n=19	PK	Res	100%	68%	21%	63%	63%	42%	16%
		Int	0%	5%	0%	0%	0%	16%	0%
		Sus	0%	26%	79%	37%	37%	42%	84%
E. cloacae n=3	U.S.A.	Res	100%	100%	100%	100%	100%	100%	100%
		Int	0%	0%	0%	0%	0%	0%	0%
		Sus	0%	0%	0%	0%	0%	0%	0%
E. cloacae n=6	PK	Res	83%	100%	67%	100%	100%	83%	67%
		Int	0%	0%	0%	0%	0%	17%	0%
		Sus	0%	0%	33%	0%	0%	0%	33%
E. aerogenes n=6	PK	Res	100%	100%	17%	100%	100%	100%	100%
		Int	0%	0%	83%	0%	0%	0%	0%
		Sus	0%	0%	0%	0%	0%	0%	0%

Species	Count.	Phenotype	CIP	SXT	GM	D*	C*
E. coli n=9	U.S.A	Res	56%	33%	22%	33%	33%
		Int	0%	0%	0%	0%	11%
		Sus	44%	67%	78%	11%	0%
E. coli n=24	PK	Res	67%	75%	38%	67%	33%
		Int	0%	0%	0%	17%	4%
		Sus	33%	25%	63%	17%	63%
K. pneumoniae n=11	U.S.A	Res	27%	36%	9%	18%	27%
		Int	0%	9%	0%	0%	0%
		Sus	73%	55%	91%	18%	9%
K. pneumoniae n=19	PK	Res	63%	68%	58%	32%	58%
		Int	11%	5%	0%	21%	11%
		Sus	26%	26%	42%	47%	32%
E. cloacae n=3	U.S.A	Res	67%	67%	33%	33%	33%
		Int	0%	0%	33%	33%	33%
		Sus	33%	33%	33%	33%	33%
E. cloacae	PK	Res	83%	83%	100%	50%	83%

n=6		Int	17%	0%	0%	17%	0%
		Sus	0%	17%	0%	33%	17%
E. aerogenes n=6	PK	Res	100%	100%	67%	83%	100%
		Int	0%	0%	0%	17%	0%
		Sus	0%	0%	33%	0%	0%

Table 2.1

Antibiotic susceptibility profiles of PH and WU clinical isolates. AM = ampicillin. CZ = cefazolin. CTT = cefotetan. CAZ = ceftazidime. CRO = ceftriaxone. FEP = cefepime. MEM = meropenem. CIP = ciprofloxacin. SXT = trimethoprim-sulfamethoxazole. GM = gentamicin. D = doxycycline. C = chloramphenicol. *For USA *E.coli* and *K. pneumoniae* in doxycycline and chloramphenicol selections, n=4. Res = Resistant, Int = intermediate, Sus= susceptible.

		Whole Genomes			
Species	**Genome**	**Number of Contigs**	**N50**	**Largest Contig**	**Total Nucleotides**
Escherichia coli	PH100	1288	35053	385222	6309191
Escherichia coli	PH101-2	404	31119	177520	4857948
Escherichia coli	PH105	288	55979	220618	5045926
Escherichia coli	PH108	754	12427	95430	4930975
Escherichia coli	PH114	280	68389	450011	5902296
Escherichia coli	PH118	1028	13550	93913	5073129
Escherichia coli	PH129	235	76630	188170	5470109
Escherichia coli	PH135	156	87914	545356	4864787
Escherichia coli	PH141	656	14806	87255	4469903
Escherichia coli	PH143	355	33581	145252	4875919
Escherichia coli	PH151-2	368	41647	143184	5066089
Escherichia coli	PH156-1	720	24256	203811	5434228
Escherichia coli	PH18	357	39599	126851	5026978
Escherichia coli	PH20	513	20802	122933	4967604
Escherichia coli	PH31	394	40363	219779	5077941
Escherichia coli	PH39	465	28063	96484	4695860
Escherichia coli	PH51	1108	13989	85055	4988648
Escherichia coli	PH5	193	185476	496282	5367128
Escherichia coli	PH85	331	45973	149681	5118463
Escherichia coli	PH90	475	31325	258054	5169559
Escherichia coli	PH92-1	738	19526	90560	5313215
Escherichia coli	PH93	487	31876	685039	5927902

Escherichia coli	PH94	325	40628	218502	4960477
Escherichia coli	PH98	401	28050	272240	4791597
Escherichia coli	WU31	502	21374	148889	4858305
Escherichia coli	WU32	329	43395	166954	5598148
Escherichia coli	WU33	487	25419	139243	4975544
Escherichia coli	WU34	278	67976	229510	5206081
Escherichia coli	WU35	267	51428	235055	4846364
Escherichia coli	WU40	145	202444	388283	5052711
Escherichia coli	WU43	193	91836	346403	5033909
Escherichia coli	WU44	196	92332	610815	5473063
Escherichia coli	WU45	332	40224	155439	4990710
Klebsiella pneumoniae	PH102	2205	3247	29455	4976237
Klebsiella pneumoniae	PH10	192	129071	506496	5877659
Klebsiella pneumoniae	PH11	224	109166	648719	5657202
Klebsiella pneumoniae	PH124	238	66778	202869	5515528
Klebsiella pneumoniae	PH12	317	46831	180775	5530414
Klebsiella pneumoniae	PH139	263	92250	433917	5458209
Klebsiella pneumoniae	PH150-2	487	40719	311295	5568393
Klebsiella pneumoniae	PH152	354	74725	275992	5863335
Klebsiella pneumoniae	PH24-1	181	209112	601209	5541053
Klebsiella pneumoniae	PH25	439	51744	186984	5437740
Klebsiella pneumoniae	PH28-1	302	72750	480318	5876774
Klebsiella pneumoniae	PH38-1	178	243482	614325	6135768
Klebsiella pneumoniae	PH40	471	43868	232550	6212797
Klebsiella pneumoniae	PH44	251	133296	479093	6297207
Klebsiella pneumoniae	PH49-2	273	84383	550379	6883299
Klebsiella pneumoniae	PH72	2195	13163	144014	7250407
Klebsiella pneumoniae	PH73	997	11652	91567	5804995

Klebsiella pneumoniae	PH88	150	289170	512144	5630043
Klebsiella pneumoniae	PH9	927	12231	72933	5207334
Klebsiella pneumoniae	WU10	179	117943	480943	5928719
Klebsiella pneumoniae	WU12	1131	12084	59345	5227362
Klebsiella pneumoniae	WU18	961	14636	96067	5664085
Klebsiella pneumoniae	WU21	372	67952	249904	5556247
Klebsiella pneumoniae	WU23	142	208533	460529	5958526
Klebsiella pneumoniae	WU2	407	55286	345910	6330941
Klebsiella pneumoniae	WU3	705	94321	408241	6853166
Klebsiella pneumoniae	WU6	297	104936	426673	5919869
Klebsiella pneumoniae	WU7	227	79691	298601	5818526
Klebsiella pneumoniae	WU8	377	64077	729491	7506385
Klebsiella pneumoniae	WU9	141	120728	292977	5557123
Enterobacter aerogenes	112-2	90	177700	566534	5153448
Enterobacter aerogenes	PH113	450	22178	97515	4695038
Enterobacter aerogenes	PH134	274	107942	399309	5126469
Enterobacter aerogenes	PH138	776	30613	144769	9901518
Enterobacter aerogenes	PH63	225	47345	184904	5211429
Enterobacter aerogenes	PH84-2	226	61644	336177	5307861
Enterobacter cloacae	PH23	179	83227	473778	5536166
Enterobacter cloacae	PH24-2	1238	20166	178617	8756501
Enterobacter cloacae	PH112-1	621	175912	517857	9155709
Enterobacter cloacae	PH125	1221	45282	215620	6004095

Enterobacter cloacae	PH158	552	34922	172567	9711730
Enterobacter cloacae	PH82	389	29307	134776	4974173
Enterobacter cloacae	WU26	538	46418	208322	5369428
Enterobacter cloacae	WU27	315	45300	155565	5251409
Enterobacter cloacae	WU29	457	24686	86433	4888311
Average	***NA***	***490***	***67666***	***285311***	***5664479***

Plasmids

Genome	Number of Contigs	N50	Largest Contig	Total Nucleotides
Acinetobacter baumannii	PH147_2	98	3072	7683	181370
Escherichia coli	WU31	41	16473	48183	214615
Escherichia coli	WU32	86	11069	31992	233185
Escherichia coli	WU33	87	9555	47149	168182
Klebsiella pneumoniae	PH11	53	19073	48703	230288
Klebsiella pneumoniae	PH88	39	15980	48463	218205
Klebsiella pneumoniae	WU13	6	12943	12943	18694
Klebsiella pneumoniae	WU14	4	12976	12976	18501
Klebsiella pneumoniae	WU17	374	4432	60538	328889
Klebsiella pneumoniae	WU18	40	18017	43330	221015
Klebsiella pneumoniae	WU19	37	22152	87768	315122
Enterobacter aerogenes	PH112_2	80	19156	52826	450788
Enterobacter aerogenes	PH113	17	5703	19905	59909
Enterobacter aerogenes	PH134	22	4168	9715	48836
Enterobacter aerogenes	PH138	171	8795	35183	466555
Enterobacter aerogenes	PH63	37	29896	44891	282470
Enterobacter cloacae	PH23	111	12334	42017	516173

Enterobacter cloacae	PH24_2	95	15201	41616	562271
Enterobacter cloacae	PH82	206	9384	37900	641720
Average	***NA***	***84***	***13178***	***38620***	***272463***

Table 2.2

Assembly metrics for whole genome and plasmid assemblies.

Chapter 3: The human microbiota as an antibiotic resistance reservoir for the *Enterobacteriaceae*

Mitchell W. Pesesky*[1], Erica C. Pehrsson*[1], Bin Wang[1,2], Janaki Lelwala-Guruge[1,2], Gautam Dantas[1,2,3]

[1]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, United States of America.

[2]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, United States of America.

[3]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri, United States of America.

*These first authors contributed equally to this work.

3.1 Abstract

The spread of antibiotic resistance threatens to undermine one of the most effective tools of modern medicine. The *Enterobacteriaceae*, in particular, are becoming resistant to the front-line treatments used against them, and in some cases have become resistant to all treatments approved for use against them. To prevent new resistance genes from entering the *Enterobacteriaceae*, we must understand which antibiotic resistance genes are accessible to them. One potential reservoir is the human microbiota, where many *Enterobacteriaceae* pathogens can reside asymptomatically, but where most of the community are members of one of two phyla other than Proteobacteria. We tested the extent to which horizontal gene transfer over long phylogenetic distances has affected resistance genes in the human microbiota by checking the codon adaptivity index of all antibiotic resistance genes from 457 intestinal bacteria sequenced as a part of the human microbiome project. We found that the difference in codon bias

between antibiotic resistance genes and their host genomes were significantly greater than between phylogenetic marker genes and their genomes, but that the distance between the two populations was small. We followed this by using functional metagenomics to find resistance genes from a 76-member mock microbiota that could provide resistance when expressed in *E. coli*. Genes that were transferred between phyla to *E. coli* were much less likely to be functional as resistance genes, though the results were gene class specific. We conclude that, with a few notable exceptions, antibiotic resistance in the intestinal microbiota is not easily accessible to *Enterobacteriaceae* pathogens.

3.2 Introduction

Antibiotic resistance is a major problem in the treatment of bacterial infections, with 23,000 deaths and up to $20 billion in direct healthcare costs attributable to resistant infections annually in the USA (1). Of primary concern are antibiotic-resistant *Enterobacteriaceae*, which reside in the intestinal environment as commensals (77) but can also cause intestinal and extraintestinal infections (27, 78, 79). Depending on their complement of antibiotic resistance genes, *Enterobacteriaceae* pathogens such as *Escherichia coli* qualify as one of the three most urgent antibiotic resistance threats in the United States according to the U.S. Centers for Disease Control and Prevention and can be resistant to all antibiotics currently used against them (1). Horizontal gene transfer (HGT) plays a major role in the spread of resistance within this family (80) and has been observed clinically between *Enterobacteriaceae* (81, 82). As the number of clinically effective antibiotics against these infections has dwindled, more focus has been placed on monitoring the spread of resistance genes to better manage the remaining treatment options.

While antibiotic-resistant pathogens are often considered in isolation, AR in non-pathogenic mixed communities that share their environment is also important for two major

reasons. First, certain mechanisms of antibiotic resistance modify or degrade antibiotics, lowering the local drug concentration and protecting proximal bacteria (including pathogens) in addition to the host. Second, non-pathogenic organisms may transfer resistance genes to pathogens through HGT. Although HGT is believed to primarily occur between closely related bacteria (83), examples of HGT between phyla have been reported (84). More distant HGT events may be more likely when the organism is under cellular stress, particularly DNA-damaging stress such as is induced by some antibiotics, facilitating the spread of resistance genes around the microbiota at the time when pathogens need them most. For example, conjugative transposons abundant in the human gut microbiome that carry resistance to tetracycline are mobilized by exposure to that antibiotic (85).

The human intestinal microbiota is an especially important potential reservoir, because it can be in contact with pathogenic *Enterobacteriaceae* during the course of infection and resulting antibiotic treatments. In most healthy humans, this community is dominated by two phyla, the Firmicutes and the Bacteroidetes (23). While that majority of these organisms are commensals, there are species from both phyla, such as *Bacteroides fragilis* and *Clostridium difficile*, which can become pathogenic under the right conditions. Previous work has shown higher levels of inter-phyla HGT in human-associated bacteria than in bacteria from other environments (19, 20). Given the already high levels of resistance in *Enterobacteriaceae*, the potential for them to acquire new resistance from distantly related bacteria in the human gut microbiota is particularly concerning.

Optimal culture media for many of the bacteria inhabiting the human gut have not been determined, so the community has primarily been characterized by genome or metagenome sequencing, but not phenotyping. A prior study used functional metagenomics to identify

antibiotic resistance genes in the microbiota (86). This technique involves creation of a DNA library by shotgun cloning a metagenome into an expression strain, which is then plated on selective media to identify genes conferring the function of interest. The study determined that the majority of functional antibiotic-resistant genes present in the human gut microbiota did not exist in public databases, suggesting that they had not yet made their way into the heavily investigated pathogenic *Enterobacteriaceae*. While sophisticated techniques exist to match DNA sequence from functional metagenomics or shotgun sequencing to the taxonomy of their most likely host bacterium (87, 88), the prevalence of HGT in the human intestinal microbiota may confuse these results. Thus, while the total resistome of the human intestinal microbiota is reasonably well understood, it is not known how HGT over long phylogenetic distances has affected the resistome in the past, and which members of the resistome have the potential to become resistance factors for pathogens in the future.

We identified antibiotic resistance genes that have likely undergone HGT from the human gut microbiota through high-quality annotation of the complete sequences of 457 publically available reference genomes spanning diverse phyla from the human gut (89). To differentiate resistance genes from the microbiota that could potentially provide resistance to pathogenic *Enterobacteriaceae*, we used functional metagenomics to model HGT from a mock human intestinal microbiota into the model *Enterobacteriaceae*, *E. coli*. We compared the genes isolated from the functional metagenomic selections to the complete genome sequences of the mock community to determine which antibiotic resistance genes were functional in *E. coli*.

3.3 Materials and Methods

Strains and Genome Sequencing: 457 proteomes of gastrointestinal tract strains were downloaded from the Human Microbiome Project (http://hmpdacc.org/catalog/, accessed March

24, 2015) for HGT analysis. 76 previously published strains with either draft or completed genomes we collected to make the mock intestinal community (MC, Table 3.1). These included 7 Actinobacteria, 19 Bacteroidetes, 11 Proteobacteria, and 39 Firmicutes. Genomic DNA was collected using phenol-chloroform extraction as follows (All liquid-handling steps were performed by a Tecan Genesis liquid handling robot unless otherwise noted): 1) Spin down 96-well, deep-well plates, containing 1 ml of culture per well at 3220 x g. 2) Aspirate supernatant. 3) Add 300 µl Buffer A (200 mM sodium chloride, 200 mM Tris, 20 mM EDTA, and 6% SDS) to each well. 4) Shake plate on orbital shaker to resuspend. 5) transfer resuspended culture to screw top tubes. 6) Add 100 µl Buffer A and 500 µl of 1 mm zirconium beads to each tube. 7) Add 500 µl Phenol:Chloroform:Isoamyl Alcohol (25:24:1, pH 7.9) to each tube. 8) Bead-beat for 5 minutes (placing tubes on ice for 2 minutes after every 2 minutes of bead beating). 9) Centrifuge for 7 minutes at 3220 x g. 10) Aspirate 500 µl from each tube to a new 96-well plate. 11) Mix 633 µl Qiagen Buffer PM and 42 µl 3M sodium acetate and add to each well of a Qiagen QIAquick 96-well plate. 12) Transfer 200 µl of sample to QIAquick plate and apply vacuum until all liquid has run through the column. 13) Repeat steps 11 and 12 until all of the sample has been applied to the columns. 14) Add 900 µl Qiagen Buffer PE to each column and vacuum until the columns are dry. 15) Repeat step 14 two times. 16) Centrifuge at 3220 x g for 7 minutes. 16) Apply vacuum for 10 minutes. 17) Centrifuge at 3220 x g for 3 minutes. 18) Add 100 µl buffer, let stand for 3 minutes. 19) Centrifuge at 3220 x g for 2 minutes to elute DNA.

For a subset of samples we did not recover enough DNA from the above robot protocol. For those samples we extracted DNA using an equivalent manual protocol in individual tubes rather than 96-well plates.

Identification of Antibiotic Resistance Genes and Phenotypic Markers: All open reading frames from the MC were identified using GeneMark (64). Antibiotic resistance genes for all genomes were identified from genome sequence using HMMER (90) and the Resfams profile Hidden Markov Model database as described previously (50). The "core" database was used, which excludes rare resistance genes with high sequence similarity to housekeeping genes that do not provide resistance. Phylogenetic markers were identified in the MC and HMP genomes using a Hidden Markov Model database of the 114 bacterial markers identified by Wu et al. 2013 (91). Where a single marker hit more than one gene in a single genome, the gene with the highest bits score was chosen.

Functional Metagenomics: Genomes were pooled in equal molar ratio based on their phylum, with a pool for Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria. A fifth pool was comprised of 22 members of the mock community expected to have high levels of resistance, based on their genome sequence, with those genomes present both in the high resistance pool and in their respective phylum pools. A sixth pool consisting of *Clostridium bolteae* alone was created as a control for high coverage of genomes in the less diverse pools. Each pool was processed as a separate sample for functional metagenomics using protocols described in detail previously (92). Briefly, we sheared genomic DNA into 2-5kb fragments, pooled the fragments from each genome together, ligated them into an expression vector, and transformed them into a susceptible *E. coli*. We then selected these *E. coli* libraries in aerobic and anaerobic conditions on 13 separate Mueller-Hinton agar plates each containing one of the following antibiotics: aztreonam (AZ), chloramphenicol (CH), cefotaxime (CT), cefoxitin (CX), ceftazidime (CZ), metronidozole (MT, anaerobic conditions only), penicillin (PE), piperacillin (PI), piperacillin-tazobactam (PITZ), tetracycline (TE), tigicycline (TG), trimethoprim (TR), and trimethoprim-

sulfamethoxazole (TRSX). We sequenced the cloned genomic fragments in the expression vector from all surviving cells on the Illumina HiSeq 2500 platform to identify the antibiotic resistance gene that permitted survival. Sequencing reads from each of these selections were aligned to each of the 76 genomes mock community genomes using Bowtie2 (62). Genes from those genomes that were fully covered by reads from the selection were computationally retrieved and the annotations of resistance genes were recorded.

Metrics and statistics: The codon adaptation index (CAI) was calculated using the formula described by Karlin et al. 1998 (22). Optimal codons for each genome were defined by determining the most frequently used codon for each amino acid across each gene in the genome. Genes that code for a high percentage of their amino acids with non-optimal codons are given a low CAI. Mean and standard error values for each distribution of CAIs were determined using the numpy python module and custom scripts. Welch's unequal variance t-test was computed using the scipy python module and custom scripts. GC content for coding regions in a genome was determined using custom scripts and significance was tested using the Fischer's exact test from the scipy python module comparing the GC content for all coding regions in a genome against the GC content of a target gene. 3-4 dinucleotide content (22) was determined by counting the occurrences of each of the 16 possible dinucleotide occurring in the third position of one codon and the first position of the following codon. Divergence of a target gene was determined by comparing all 3-4 dinucleotides from that gene against all genes in that genome using the Fischer's exact test as encoded by the scipy python module.

3.4 Results

To determine if antibiotic resistance genes were transferring over large phylogenetic distances, we measured the codon adaptation index (CAI) (22) over all antibiotic resistance

genes and phylogenetic markers from completed genomes from the Human Microbiome Project (HMP, Figure 3.1a). CAI assesses the codon bias of a particular gene relative to the genome wide codon usage, with those genes showing strong biases potentially being horizontally transferred. The distributions for both antibiotic resistance genes and phylogenetic markers are approximately normal, though each have an extended lower tail. Though the genomes largely overlap and have similar means, 0.719 for resistance genes and 0.738 for phylogenetic markers, their difference is statistically significant (Welch's unequal variances t-test: p=8.23e-84, 95% CI of the difference: 0.018-0.022). This difference is potentially due a larger proportion of samples in the lower tail for the resistance genes (Figure 3.1a), lowering the population mean.

We next desired to test if resistance genes from the microbiota could be transferred over large phylogenetic distances in the future, and if this corresponded to whether or not they had undergone HGT in the past. To this end we assembled a mock human intestinal community (MC) from 76 strains (Table 3.1) and performed the same CAI analysis (Figure 3.1b). The CAI distributions showed similar properties for the MC as it had for the HMP genomes, including largely overlapping distributions, long lower tails, and similar but significantly distinct population means (resistance gene mean: 0.728, marker gene mean: 0.742, Welch's unequal variances t-test: 5.31e-09, 95% CI of the difference: 0.010-0.019).

We next used functional metagenomics to assess which genes from the MC had the potential to be used by *E. coli* as resistance factors (Figure 3.2). Genes are reported at the most specific family to which they could confidently be assigned, meaning that some families are subsets of others, such as the TEM (RF0126) and Class A (RF0053) β-lactamases. Because functional metagenomics relies on shotgun cloning, stochastic factors can prevent some genes

that are present in the source material from being expressed in the library. As a positive control for this we looked at TR resistance, which can be caused by overexpression of the TR target, dihydrofolate reductase (DHFR). We would expect, then, that each DHFR encoded in the source genomes should be able to serve as a resistance gene in our libraries, and since each organism should encode at least one DHFR, we can use this as a proxy for our total library coverage. In the few cases where a DHFR could not be identified in the genome sequence, then it would also not be detected by functional metagenomics, since we interpreted or functional metagenomic results by aligning reads to the reference genomes. From the Actinobacteria, Bacteroidetes, and Proteobacteria, we detected almost all of the DHFR genes (RF0063), 83%, 89%, and 90%, respectively. Unfortunately, for the Firmicutes we only recovered 56% of the DHFR genes, suggesting that there was less complete coverage of this phylum in our libraries.

In general, we noted that a large number of the resistance genes originated from the Proteobacteria, though made up only 14% of the total MC. This was as expected, since Proteobacterial genes are more likely to be able to integrate into the transcriptional and metabolic network of the *E. coli* host than genes from other phyla, particularly in the case of regulatory genes such as *marA* (RF0091), *marR* (RF0092), and *ramA* (RF0114). Interestingly, AraC transcription factors (RF0079) that were not part of the MarA or RamA families were not found to confer resistance, except for one from the Firmicute *Dorea longicatena*.

One of the resistance protein families that was both highly prevalent in the MC, and readily used by *E.coli* was the TetW-like ribosomal protection proteins (RF0133). Though these proteins originated largely from the Firmicutes and one Actinobacteria, they were almost universally captured by our functional metagenomic selection.

Unlike TetW, the Class A β-lactamases from the Firmicutes were generally not captured by functional metagenomics. Only 13% of the Firmicute Class A β-lactamases were expressed in *E. coli*, even when including the TEM subclass. To determine if the two captured Firmicute β-lactamases were more similar to Proteobacterial Class A β-lactamases, than the non-captured enzymes from their phylum, we performed a multiple sequence alignment of all complete Class A β-lactamases from both phyla (several protein sequences were truncated in their reference sequence). We then generated a protein phylogeny from this alignment of 18 sequences (Figure 3.3). The four Class A β-lactamases that were captured by functional metagenomics cluster in one branch of the tree, with the TEM from *Clostridium nexile* showing >99% nucleotide sequence identity to the TEM from *Escherichia fergusonii*.

Since two Firmicute Class A β-lactamases showed potential for inter-phyla exchange, we next tested if they showed signs of evolving in a different genomic context, using three measures of genome adaptation: GC content, 3-4 dinucleotide usage, and the CAI. The gene from *C. nexile* had a GC content of 49.12%, compared to the genome wide GC content of 39.02% (Fischer's exact p-value = 2.0e-9), a significantly divergent dinucleotide usage (Fischer's exact p-value = 9.2e-4), and a CAI of 0.602, which is in the lower tail of resistance genes from the MC (Figure 3.1b). In contrast, the GC content of the nearly identical gene in *E. fergusonii* (49.01%) was much closer to the genome-wide GC content of 49.93% (Fischer's exact p-value = 0.61), and had more similar dinucleotide usage (Fischer's exact p-value = 0.12), though the CAI was similarly low at 0.622. Finally the Class A β-lactamase from *Clostridum bolteae* was the only Class A β-lactamase to be expressed in *E. coli* that was not a part of either the TEM or CblA (from Bacteroidetes) known subclasses. Its GC content was 30.87% compared to the genome wide GC

content in *C. bolteae* of 49.05% (Fischer's exact p-value = 2.2e-16), and a divergent dinucleotide usage (Fischer's exact p-value ≤ 1e-6), though the CAI of 0.640 was closer to the average for resistance genes.

3.5 Discussion

In looking at both the MC and HMP genomes we find a significant, but small enrichment for probable HGT events in resistance genes compared to phylogenetic marker genes. Since genes in the marker gene set were chosen in part based on how closely a phylogeny based on them agreed with the species phylogeny (91), it is likely that they represent the minimal HGT that can be expected over a gene set of that size. If we had compared resistance genes against another gene set, such as accessory metabolic genes, then we may not have seen a significant difference in likely HGT events. The CAI, though, may underestimate HGT events for two reasons. First, organisms that are closely related to each other are likely to have similar codon biases, so the CAI is likely to remain high for genes that have been transferred between closely related bacteria. We were focused on transfer events between the *Enterobacteriaceae* and other members of the intestinal microbiota, most of which are from different phyla, so we expect that this limitation of CAI had only minor effects on our analysis. Second, there is a finite set of primary codons that an organism can use to code for its amino acids, and therefore it is possible that two bacteria with very different evolutionary histories could end up having similar codon biases. This is one possible explanation for why the Class A β-lactamase in *C. bolteae* has a relatively average CAI, though the gene looks like it has come from a very different genome by other metrics. We also expect this limitation to have had a minimal effect, though, because there are over 6.2e15 possible combinations of primary codons, and so convergent evolution of codon biases is likely to be a rare event.

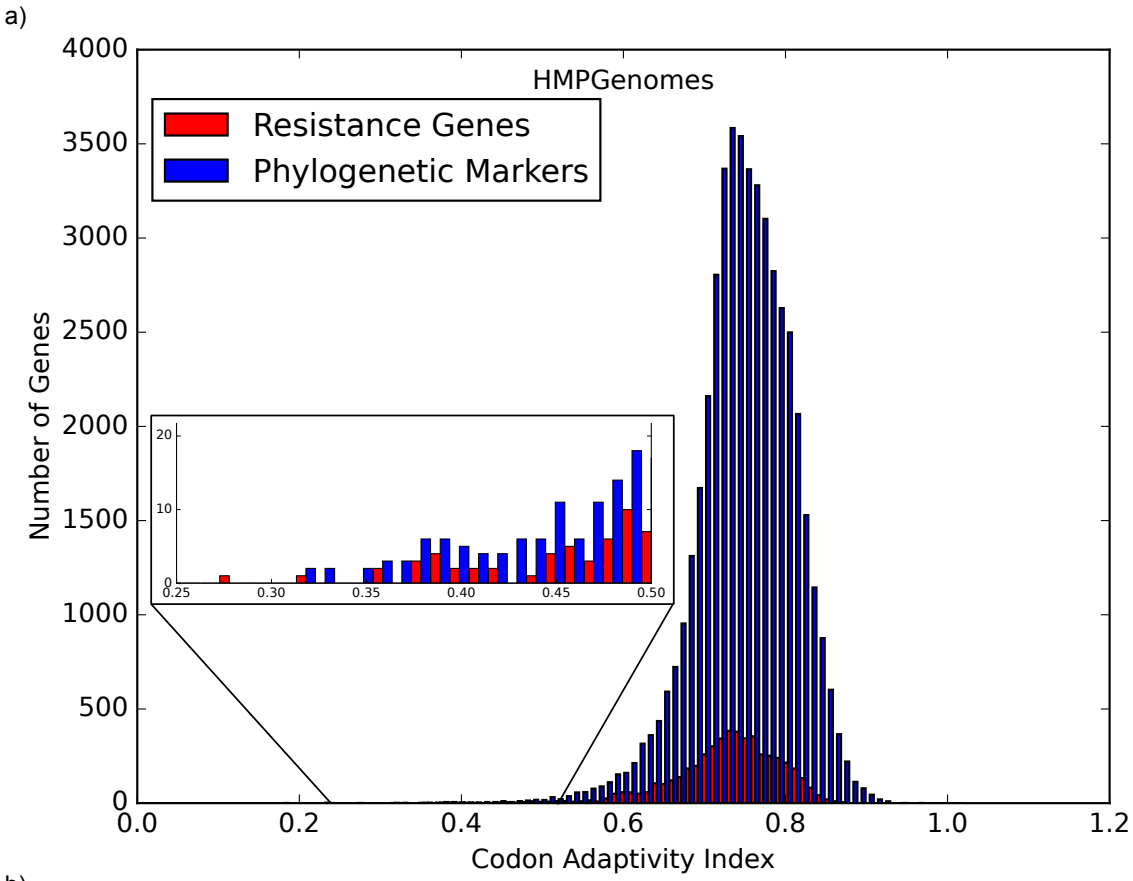
Much like our CAI analysis showed that inter-phyla HGT has had little impact on the human microbiota in the past, our functional metagenomic analysis demonstrated that inter-phyla HGT is unlikely to have a major impact on the *Enterobacteriaceae* resistome in the future. With a few notable exceptions, the potential resistance genes identified in the Actinobacteria, Bacteroidetes, and Firmicutes did not provide resistance to *E. coli*. It is likely that some of the predicted resistance genes from these groups are not true resistance genes, particularly the ABC transporters which are a known issue in the Resfams database (50); however, Resfams has proven to be highly reliable on resistomes from diverse environments and phylogenies. Given that all of the bacteria in the microbiota have been put under increased selection pressure by the medical use of antibiotics, we expect that most of the resistance genes detected in these genomes are true resistance genes, but they cannot be used by *E. coli* for resistance.

The Class A β-lactamases are an interesting case because it seems that they are unlikely to be effective if transferred from Firmicutes to the *Enterobacteriaceae*. One could expect that enzymes that act directly on the antibiotic would be effective regardless of genetic context, especially if they were expressed from a native promoter, as was the case in our functional metagenomic assay. The most likely explanation is that the export signals for β-lactamases differ between the phyla. If this is the case, then the Firmicute Class A β-lactamases are active, but trapped in the *E. coli* cytoplasm rather than the periplasm. Since these enzymes degrade β-lactams though, a large number of Firmicute β-lactamases in the human microbiota may still provide protection to susceptible *Enterobacteriaceae*.

Overall, we show that although many antibiotic resistance genes exist in the human microbiota environment, it is unlikely that many of them will contribute to the

Enterobacteriaceae resistome. Some *Enterobacteriaceae* are already resistant to all known antibiotics (1), so the inaccessibility of this reservoir may seem of little comfort for now. As new antibiotics are developed, though, this opens the possibility of using narrow spectrum treatments against the *Enterobacteriaceae* even when resistance genes are prevalent in other phyla. The principle of low inter-phyla HGT should also be tested using models for other major multi-drug resistant pathogens as hosts for diverse resistance genes. If it holds true then the human microbiota and other environments would not be single, highly-connected resistance reservoirs, but collections of smaller, loosely connected or unconnected reservoirs, each inaccessible to the other's pathogens.

3.6 Figures


a)
HMPGenomes
Resistance Genes
Phylogenetic Markers
Number of Genes
Codon Adaptivity Index
0
500
1000
1500
2000
2500
3000
3500
4000
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
10
20
0.25
0.30
0.35
0.40
0.45
0.50

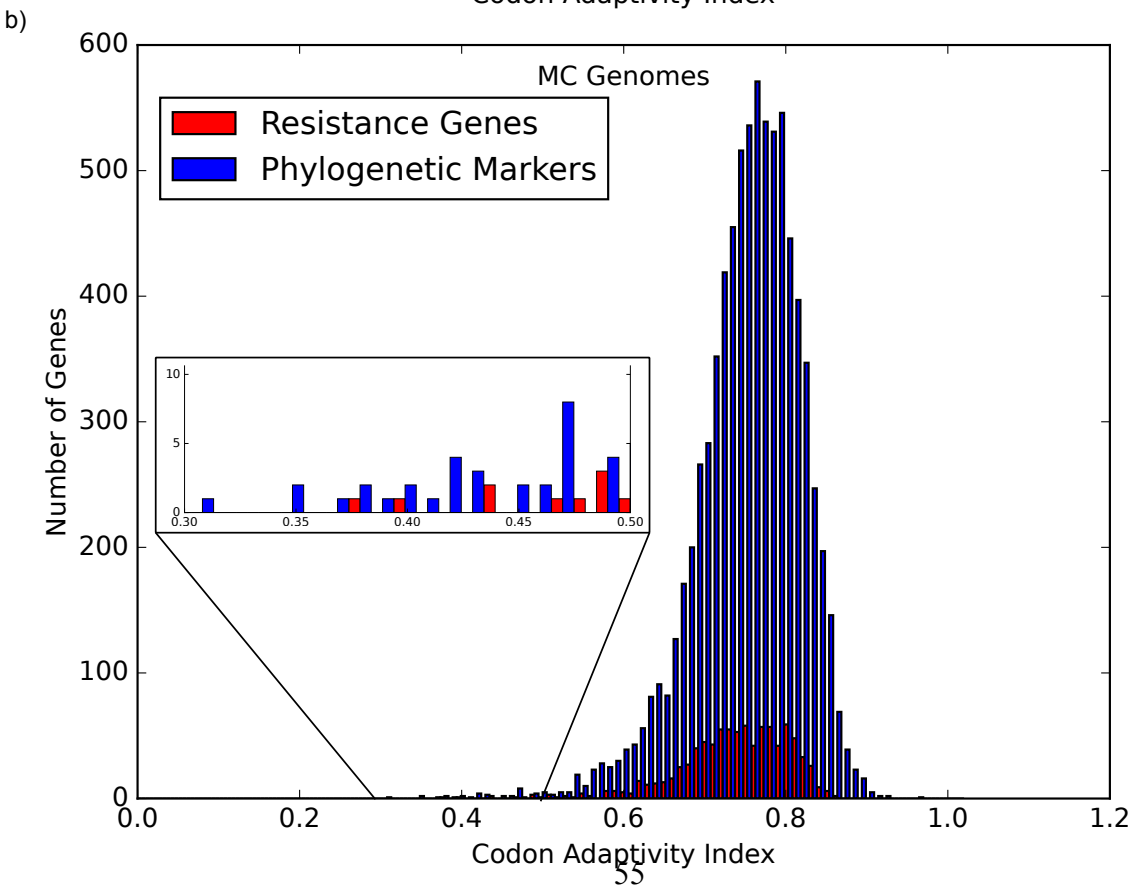
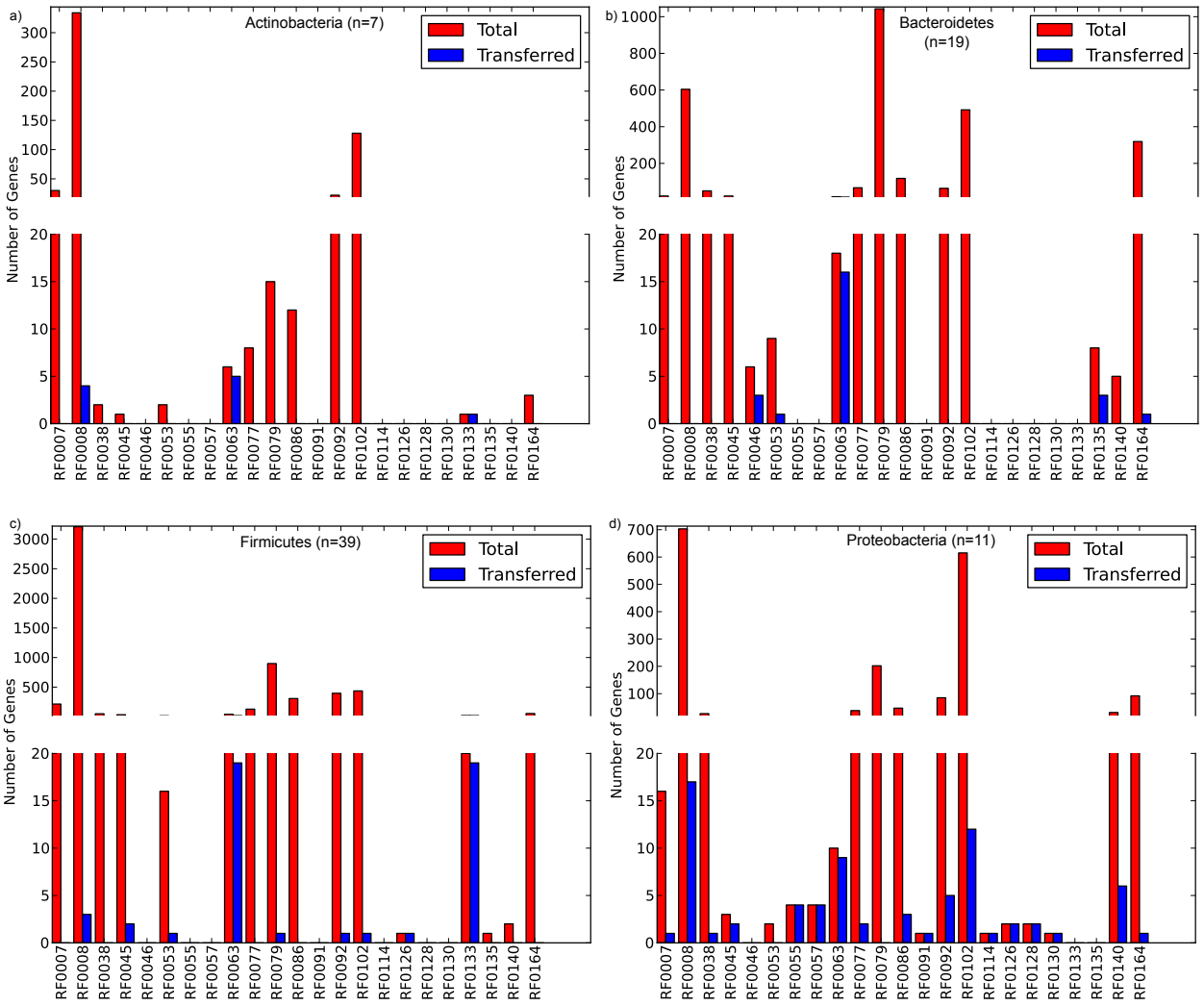

b)
MC Genomes
Resistance Genes
Phylogenetic Markers
Number of Genes
Codon Adaptivity Index
0
100
200
300
400
500
600
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
5
10
0.30
0.35
0.40
0.45
0.50

Figure 3.1

Distribution of codon adaptation indices for each phylogenetic marker gene and antibiotic resistance gene across all genomes in the a) HMP genomes or b) MC genomes. Insets show the lower tail of each distribution.



e)

Resfam ID	Protein Family	Resistance
RF0086	β-lactamase superfamily	PE, PI, CX
RF0091	marA Transcription Factor	PE ,CH, TE
RF0092	MarR	PE, PI, PITZ, CH, TE
RF0102	MFS Transporter	PE, PI, CH, TE
RF0114	ramA	PE, PI, CH, TE
RF0126	TEM β-lactamase	β-lactams
RF0128	TetA(B) Efflux pump	TE
RF0130	TetD efflux pump	TE
RF0133	TetW Ribosomal Protection Protein	TE
RF0135	Ribosomal Protection Protein	TE
RF0140	Bcr-CflA Efflux pump	TE
RF0164	RND_mfp	AZ, CZ, PI

Resfam ID	Protein Family	Resistance
RF0007	ABC Efflux	TE
RF0008	ABC Transporter	TE, MT
RF0038	β-lactamase	β-lactams
RF0045	Chloramphenicol Acetyltransferase	CH
RF0046	CblA β-lactamase	β-lactams
RF0053	Class A β-lactamase	β-lactams
RF0055	Class C β-lactamase	β-lactams
RF0057	CMY β-lactamase	β-lactams
RF0063	Dihydrofolate Reductase	TR, TRSX
RF0077	Glyoxalase	PI, PITZ
RF0079	AraC Transcription Factor	TE

Figure 3.2

Functionality of antibiotic resistance genes from four phyla in an *E. coli* background. Resistance genes belonging to one of the 23 families functional in *E. coli* in this experiment were identified either from the source genomes (red) or from the functional metagenomic selections (blue). Genes were derived from strains in the MC from a) the Actinobacteria, b) the Bacteroidetes, c) the Firmicutes, or d) the Proteobacteria. Descriptions of each protein family and the antibiotics resisted by genes in that family in this experiment are shown in e. Some specific genes provided only a subset of the total resistance seen for their family.

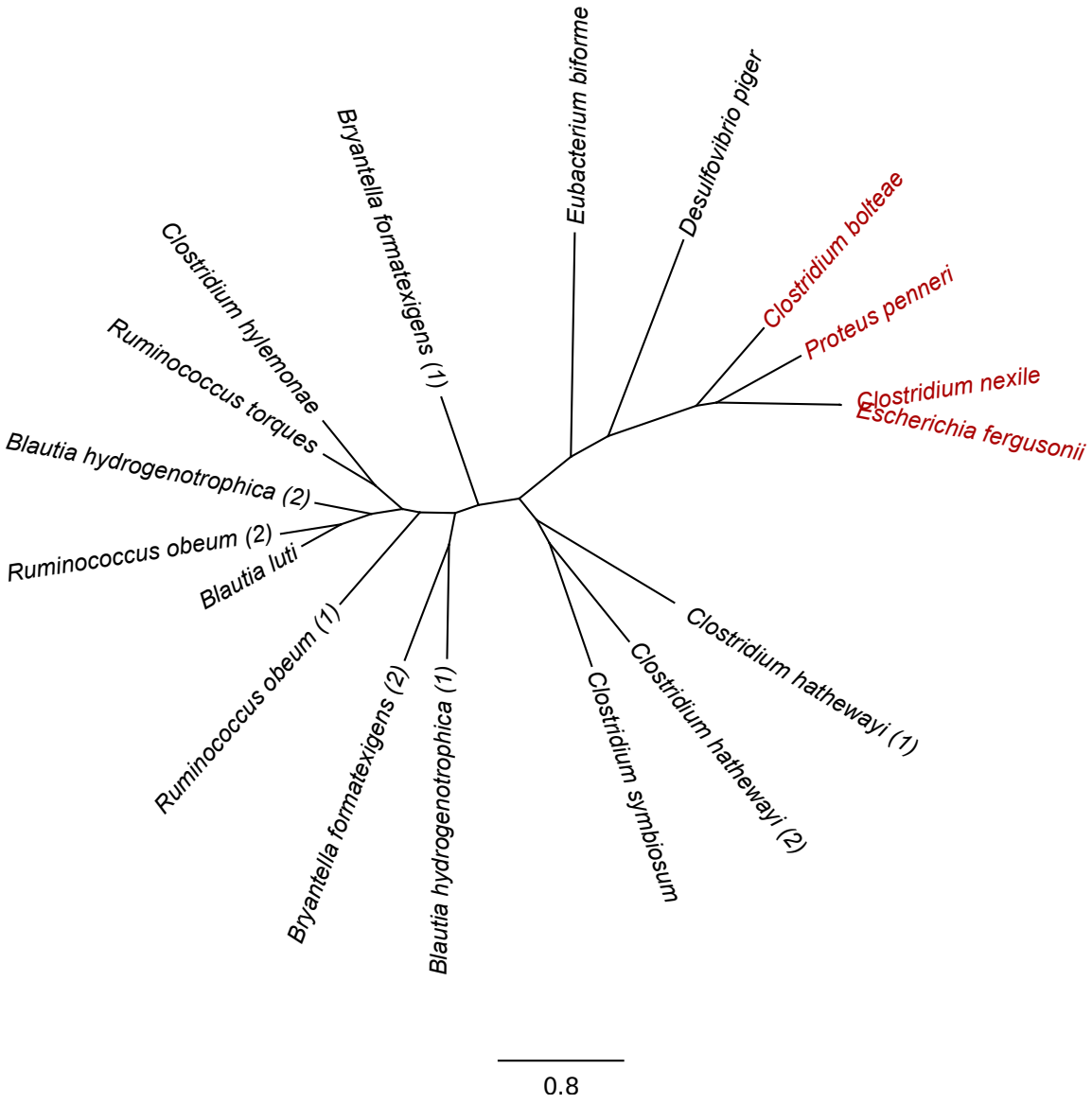


Figure 3.3

Protein tree for all complete Class A β-lactamases from the Firmicutes and Proteobacteria, labeled with the species of origin. Red species names indicate proteins that provided resistance to *E. coli* in the functional metagenomic experiment. Where more than one Class A β-lactamase was found in the same genome, the sequences are distinguished by a number in parenthesis. Scale bar shows substitutions per site.

3.7 Tables

Phylum	Genus	Species	Strain ID	NCBI GI	Libraries
Bacteroidetes	Alistipes	indistinctus	DSM 22520	354605450	Bact
Firmicutes	Anaerococcus	hydrogenalis	DSM 7454	207998074	Firm
Firmicutes	Anaerotruncus	colihominis	DSM 17241	163805931	Firm, Res
Bacteroidetes	Bacteroides	caccae	ATCC 43185	134302814	Bact
Bacteroidetes	Bacteroides	cellulosilyticus	DSM 14838	221216683	Bact
Bacteroidetes	Bacteroides	cellulosilyticus	WH2	661525291	Bact
Bacteroidetes	Bacteroides	coprophilus	DSM 18228	221217348	Bact
Bacteroidetes	Bacteroides	dorei	DSM 17855	208431349	Bact
Bacteroidetes	Bacteroides	eggerthii	DSM 20697	208431420	Bact
Bacteroidetes	Bacteroides	finegoldii	DSM 17565	239514184	Bact
Bacteroidetes	Bacteroides	intestinalis	DSM 17393	172085037	Bact
Bacteroidetes	Bacteroides	ovatus	ATCC 8483	145225849	Bact
Bacteroidetes	Bacteroides	stercoris	ATCC 43183	163805998	Bact
Bacteroidetes	Bacteroides	thetaiotaomicron	VPI 5482	29345410	Bact
Bacteroidetes	Bacteroides	uniformis	ATCC 8492	153860072	Bact, Res
Bacteroidetes	Bacteroides	vulgatus	ATCC 8482	150002608	Bact
Bacteroidetes	Bacteroides	xylanisolvens	DSM 18836	479162165	Act
Actinobacteria	Bifidobacterium	adolescentis	L2-32	145845910	Act
Actinobacteria	Bifidobacterium	bifidum	DSM 20456	705432742	Act
Actinobacteria	Bifidobacterium	dentium	ATCC 27678	169822569	Act
Actinobacteria	Bifidobacterium	pseudocatenulatum	DSM 20438	222444389	Act
Firmicutes	Blautia	hansenii	DSM 20583	255747248	Firm, Res
Firmicutes	Catenibacterium	mitsuokai	DSM 15897	218294255	Firm
Proteobacteria	Citrobacter	youngae	ATCC 29220	282597575	Prot, Res

Firmicutes	Clostridium	asparagiforme	DSM 15981	223982486	Firm
Firmicutes	Clostridium	bartlettii	DSM 16795	163813846	Firm
Firmicutes	Clostridium	bolteae	ATCC BAA-613	160942467	Firm, Res, Bol
Firmicutes	Clostridium	hathewayi	DSM 13479	229504944	Firm
Firmicutes	Clostridium	hiranonis	DSM 13275	207998176	Firm
Firmicutes	Clostridium	hylemonae	DSM 15053	224228175	Firm
Firmicutes	Clostridium	nexile	DSM 1787	210619843	Firm, Res
Firmicutes	Clostridium	ramosum	DSM 1402	163806084	Firm
Firmicutes	Clostridium	scindens	ATCC 35704	163806067	Firm, Res
Firmicutes	Clostridium	sp. M62/1		281423074	Firm, Res
Firmicutes	Clostridium	sp. SS2/1		163805974	Firm, Res
Firmicutes	Clostridium	spiroforme	DSM 1552	167642790	Firm
Firmicutes	Clostridium	sporogenes	ATCC 15579	183569454	Firm
Firmicutes	Clostridium	symbiosum	ATCC 14940	545400804	Firm
Actinobacteria	Collinsella	aerofaciens	ATCC 25986	146334937	Act, Res
Actinobacteria	Collinsella	intestinalis	DSM 13280	229816750	Act
Actinobacteria	Collinsella	stercoris	DSM 13279	208430894	Act
Firmicutes	Coprococcus	comes	ATCC 27758	224983101	Firm, Res
Firmicutes	Coprococcus	eutactus	ATCC 27759	163816861	Firm, Res
Proteobacteria	Desulfovibrio	piger	DSM 749	209954086	Prot
Firmicutes	Dorea	formicigenerans	ATCC 27755	163816878	Firm, Res
Firmicutes	Dorea	longicatena	DSM 13814	138276246	Firm, Res
Proteobacteria	Edwardsiella	tarda	ATCC 23685	284798543	Prot
Proteobacteria	Enterobacter	cancerogenus	ATCC 35316	260842206	Prot
Proteobacteria	Escherichia	coli	K12 MG1655	556503834	Prot, Res
Proteobacteria	Escherichia	fergusonii	DSM	218547440	Prot, Res

			13698		
Firmicutes	Eubacterium	biforme	DSM 3989	210137602	Firm
Firmicutes	Eubacterium	dolichum	DSM 3991	160916359	Firm, Res
Firmicutes	Eubacterium	eligens	ATCC 27750	238915976	Firm
Firmicutes	Eubacterium	rectale	ATCC 33656	238922432	Firm
Firmicutes	Eubacterium	ventriosum	ATCC 27560	145197224	Firm
Firmicutes	Faecalibacterium	prausnitzii	M21/2	160946040	Firm
Firmicutes	Holdemania	filiformis	DSM 12042	223986999	Firm, Res
Firmicutes	Lactobacillus	reuteri	DSM 20016	148543243	Firm
Firmicutes	Marvinbryantia	formatexigens	DSM 14469	241911894	Firm
Firmicutes	Megamonas	funiformis	DSM 19343	375087558	Firm, Res
Firmicutes	Mitsuokella	multacida	DSM 20544	253946801	Firm, Res
Bacteroidetes	Parabacteroides	distasonis	ATCC 8503	150006674	Bact
Bacteroidetes	Parabacteroides	johnsonii	DSM 18315	209953680	Bact
Bacteroidetes	Parabacteroides	merdae	ATCC 43184	145218793	Bact
Proteobacteria	Proteus	penneri	ATCC 35198	224959514	Prot
Proteobacteria	Providencia	alcalifaciens	DSM 30120	209953830	Prot
Proteobacteria	Providencia	rettgeri	DSM 1131	264671240	Prot
Proteobacteria	Providencia	stuartii	ATCC 25827	172072827	Prot
Firmicutes	Roseburia	intestinalis	DSM 14610	239936509	Firm
Firmicutes	Ruminococcus	gnavus	ATCC 29149	146386260	Firm, Res
Firmicutes	Ruminococcus	hydrogenotrophicus	DSM 10507	223987233	Firm, Res
Firmicutes	Ruminococcus	lactaris	ATCC 29176	197304080	Firm
Firmicutes	Ruminococcus	obeum	ATCC 29174	138312116	Firm, Res
Firmicutes	Ruminococcus	torques	ATCC 27756	138263299	Firm
Firmicutes	Streptococcus	infantarius	ATCC BAA-	170764697	Firm

			102		
Firmicutes	Subdoligranulum	variabile	DSM 15176	260590249	Firm, Res

Table 3.1

Strains comprising the mock intestinal community (MC). Final column indicates the functional metagenomic libraries that strain was included in. Act = Actinobacteria library, Bact = Bacteroidetes library, Firm = Firmicutes library, Prot = Proteobacteria library, Res = Expected resistance library, Bol = *Clostridium bolteae* genomic library.

Chapter 4: Evaluation of Machine Learning and Rules-Based approaches for predicting antimicrobial resistance profiles from whole genome sequence data

Mitchell W Pesesky[1], Tahir Hussain[1,2], Meghan Wallace[3], Sanket Patel[1,3], Saadia Andleeb[2], Carey-Ann D. Burnham[3,4], Gautam Dantas[1,3,5]

[1]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri, United States of America

[2]Atta ur Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad, Pakistan

[3]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, United States of America

[4]Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri, United States of America

[5]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri, United States of America

4.1 Abstract

Judicious use of antimicrobial agents is a global public health priority. The time-to-result for culture-based microorganism recovery and phenotypic antimicrobial susceptibility testing necessitate initial use of empiric (frequently broad-spectrum) antimicrobial therapy. New sequencing technologies with rapid result output are emerging; thus, direct, whole genome sequencing of microbes is one approach to reducing the time to generate antimicrobial resistance

information. We evaluated two algorithms for automated translation of bacterial whole genome sequence data into isolate-specific susceptibility profiles on 78 clinical *Enterobacteriaceae* isolates. The first, a rules-based algorithm, makes predictions using current knowledge of the characteristics of resistance genes in the *Enterobacteriaceae*. The second was a machine learning algorithm that predicted resistance and susceptibility based on a list of annotated resistance factors. The rules based and machine learning predictions achieved agreement of 86.4% and 91.0%, respectively, with phenotypic antimicrobial susceptibility testing for 12 antimicrobial agents. Novel variants of known resistance factors, incomplete genome assembly, and variability in the phenotypic antimicrobial susceptibility testing contributed to prediction errors. We also employed machine learning to make quantitative predictions about the level of resistance of each isolate, demonstrating a potential advantage of the machine learning approach over rules-based or manual interpretation of resistance data. Sequence-based antimicrobial susceptibility testing shows great promise as a diagnostic tool, and we outline specific research goals to further refine the utility of this methodology.

4.2 Introduction

The spread of antibiotic resistance has become an urgent threat to modern medicine's control over bacterial infections. In critically ill patients, it has been well established that the time taken to administer an appropriate antibiotic agent inversely correlates with improved patient outcomes (49). Unfortunately, definitive *in vitro* antibiotic susceptibility testing (AST) results are not typically available until at least two days after specimens arrive in the clinical laboratory (48), necessitating broad spectrum empiric antibiotic therapy.

New diagnostic methods that reduce the interval of diagnostic uncertainty could reduce the time to optimization of antibiotic therapy for the millions of patients infected with antibiotic

resistant pathogens each year (1), and potentially reduce some of the pressure causing the increase in the prevalence of antibiotic resistance. Several diagnostic assays are emerging that rapidly identify antibiotic resistance based on genotypic rather than phenotypic information, including multiplex PCR and microarray assays designed to identify resistance-specific markers (93-95). While these techniques can be successful at detecting a limited subset of resistance determinants mediated by specific enzymes, few of these methods claim detection of resistance mediated by target mutation, such as fluoroquinolone resistance (93-95) and accuracy decreases as additional genes are assayed (96). Whole genome sequencing (WGS) has been proposed as an alternative method for identifying antibiotic resistance genes (48, 97). As this approach interrogates the entire genome of each organism, WGS can also be used to identify novel resistance genes and target-mediated resistance. Another advantage is that neither the DNA preparation nor the resistance gene identification steps increase in duration as the number of antibiotics evaluated increases. Pathogen WGS is also being evaluated for additional clinical tasks, such as species identification (98) and assessment of strain relatedness (99, 100), allowing a single test to be used for multiple purposes. Several authors have implemented WGS-based AST (101-104), arguing that the time and cost of sequencing will decrease to acceptable levels for clinical diagnostics in the near future. Despite its potential upside, WGS-based AST still has several hurdles to overcome before it can become a viable alternative to *in vitro* phenotypic tests. First, it is currently slower and more expensive than other technologies. This hurdle is likely to diminish over time, as innovations in sequencing methodologies are continually decreasing cost and assay duration (105, 106). Second, like with PCR or microarray-based assays, additional analysis is required to translate gene identification into expected susceptibility profiles. Previous reports of WGS-based AST have focused on identifying antibiotic resistance determinants,

without optimizing interpretation of the organism susceptibility profile. We have generated an automated pipeline for predicting antibiotic susceptibility from full genome sequence, including identification of resistance determinants and susceptibility profile determination. For the critical step of translating resistance gene identifications into phenotype predictions, we compared a rules-based algorithm to a machine learning approach. We applied this pipeline to predict the resistance profiles of 78 genome sequenced *Enterobacteriaceae* isolates (80) to 12 antibiotics, and compared the results to the categorical interpretation of *in vitro* susceptibility testing determined by Kirby Bauer Disk Diffusion (12) for the same isolates and antibiotics. For each antibiotic class, we then determined if the cause for errors in our algorithm could be identified from manual analysis of the genomic data. Our goal was to determine the most effective implementation of automated interpretation for WGS-based AST, and to identify specific research objectives that could improve future versions.

4.3 Methods and Materials

Sample Selection, Phenotype Determination, and Sequencing. Isolates were retrieved from existing strain banks in Pakistan Railway General Hospital, Rawalpindi, Pakistan; the Pakistan Institute of Medical Sciences in Islamabad, Pakistan; or Barnes Jewish Hospital/Washington University (WU) School of Medicine in Saint Louis, Missouri, U.S.A. Isolate phenotypes and draft genome sequences were determined as described in detail previously (80). Briefly, antibiotic susceptibility phenotypes were determined using Kirby-Bauer disk diffusion according to Clinical and Laboratory Standards Institute guidelines and interpretive criteria (12). Each isolate was sequenced on the Illumina Hi Seq 2500 using 101 bp reads, and reads were assembled into draft genomes using Velvet (63). Genes were annotated using HMMER3 (65) comparisons with the Pfams (107), TIGRFAMs (108), and Resfams (50) databases.

Antibiotic Resistance Prediction. Antibiotic susceptibility was predicted for each isolate using a rules-based (RB) algorithm and a Logistic Regression (LR) algorithm with Resfams annotated genes as the inputs. The *in vitro* phenotypic susceptibility results were used as the gold standard for comparison with our genotype-based AST. Errors in the genotype-based AST were defined in relation to the *in vitro* results, with major errors being a resistant prediction from the WGS-based AST being discrepant with an *in vitro* susceptible phenotype and very major errors being a susceptible prediction from the WGS-based AST being discrepant with an *in vitro* resistant phenotype. For the purposes of this comparison, intermediate resistance phenotypes from the phenotypic AST were counted as resistant. The RB algorithm was built using custom python scripts, while the LR algorithm used Weka (109) with the following parameters: (weka.classifiers.functions.Logistic -R 1.0E-8 -M 10). The prediction model built by logistic regression was evaluated by leave-one-out cross validation (110).

β-lactam antibiotics. For the RB algorithm, we identified each specific β-lactamase in a given isolate, and matched that identity to the set of β-lactam antibiotics to which it was expected to give resistance (www.lahey.org/Studies, accessed March 25th, 2014 (70)). For beta-lactamases that were identified that had less than 100% amino acid identity to a known beta-lactamase, we assigned it to the resistance profile equal to the most resistant member of its gene family (i.e. CTX-M, KPC, etc.). The expected resistance profile for the isolate was the union of those individual β-lactamase sets, and the predicted susceptibility profile was the inverse of the resistance profile. For the LR algorithm we used isolate species and the presence or absence of specific β-lactamase families as input.

Ciprofloxacin. We predicted susceptibility to ciprofloxacin by comparison of the quinolone resistance determining regions (QRDR) of *gyrA* (residues 68-106), *parC* (residues 68-106), and

gyrB (residue 426) for each isolate against the wild type. For the RB algorithm, ciprofloxacin susceptibility was predicted for an isolate only if it had the wild-type QRDR in both genes. If the *gyrA* or *parC* genes were not completely assembled and no definite resistance mutation could be identified, than the resistance prediction was not determined (N.D.). N.D. resistance was not counted as a major error or very major error, but was counted against the percent accuracy of the prediction. We also determined the presence or absence of *qnr* genes and quinolone efflux transporter analogues (*oqxAB*), to identify their effect on our predictive accuracy. For the LR algorithm, isolate species and variations in the QRDR were used as inputs.

Doxycycline and Chloramphenicol. For the RB algorithm we predicted isolates to be susceptible to doxycycline only in the absence of any genes with known tetracycline resistance phenotypes (ex. *tetA*). The same metric was used for chloramphenicol, where examples of chloramphenicol resistance genes include chloramphenicol acetyl-transferases and chloramphenicol efflux pumps. The LR algorithm for both antibiotic conditions used isolate species and resistance gene family identity as inputs.

Gentamicin. For the RB algorithm we determined susceptibility to gentamicin by comparing the sequence of identified aminoglycoside resistance genes against a database of known profiles, in this case CARD (111). If any of an isolate's aminoglycoside resistance genes had previously been reported to provide resistance to gentamicin, we predicted the isolate to be resistant, otherwise we predicted it to be susceptible. For the LR algorithm we used isolate species and presence or absence of each resistance gene family identity as inputs.

Trimethoprim-sulfamethoxazole. For the RB algorithm, we predicted susceptibility to trimethoprim-sulfamethoxazole by enumerating the unique dihydrofolate-reductase (DHFR)

enzymes present within each isolate, defining unique as having less than 95% nucleotide identity to any other DHFR in the genome. If two divergent DHFRs were present within the same genome, that isolate was predicted to be resistant, otherwise it was predicted to be susceptible. For the LR algorithm we clustered all of the DHFR genes at 95% nucleotide identity using cd-hit (112) generating 20 clusters. We used isolate species and the presence or absence of each DHFR cluster as inputs.

Zone of inhibition prediction. The size of zones of inhibition for each isolate were predicted using the random forests algorithm via Weka, using the parameters: (weka.classifiers.trees.RandomForest -I 100 -K 0 -S 1 -num-slots 1). Inputs were the same as for the LR algorithm except that the class variable was size of the zone of inhibition, rather than the "Susceptible" or "Resistant" label. As with the LR algorithm, predictions models were evaluated by leave-one-out cross validation.

Functional Metagenomics. For functional metagenomics (113) we utilized protocols detailed previously (114). Briefly, we sheared genomic DNA into 2-5kb fragments, pooled the fragments from each genome together, ligated them into an expression vector (pZE21), and transformed them into a susceptible *E. coli*. We then treated these recombinant *E. coli* libraries with penicillin, amoxicillin, cefotaxime, ceftazidime, meropenem, aztreonam, gentamicin, tetracycline, tigecycline, chloramphenicol, and trimethoprim in separate solid Mueller-Hinton agar selection experiments. We sequenced the cloned genomic fragments in the expression vector from all surviving cells on the Illumina HiSeq 2500 platform to identify the antibiotic resistance gene that permitted survival. Sequencing reads from each of these selections were aligned to each of the 78 genomes using Bowtie2 (62). Genes from those genomes that were

fully covered by reads from the selection were computationally retrieved and annotated using the Pfams (107), TIGRFAMs (108), and Resfams (50) databases.

4.4 Results

Prediction algorithm performance: The first step in our resistance prediction algorithm is highly accurate resistance gene identification using the Resfams database (50). Many of the isolates in this set were multidrug resistant by the phenotypic AST, and several were resistant to all antibiotics tested (Table 4.1). Though the average number of resistance genes varied between the species in this set, resistance gene families were shared between species (Figure 4.1). We next applied LR machine learning and RB algorithms to predict antibiotic susceptibility from these gene identifications, using their *in vitro* phenotypes as the gold standard (Figure 4.2). Our RB algorithm had overall agreement of 86.6% with the phenotypic AST, with an overall major error rate of 6.9% and very major error rate of 4.8%. Agreement for individual antibiotics ranged from 79.5% to 96.2% while major error rates ranged from 0% to 18% and very major error rates ranged from 0% to 12.8% (Figure 4.S1). The LR algorithm had a higher overall agreement of 91.0% to the phenotypic ASTs, with a major error rate of 2.4% and a very major error rate of 6.6% (Figure 4.1). Individual antibiotic agreement ranged from 80.8% to 97.4%, while major error rates ranged from 0% to 6.4% and very major error rates ranged from 1.3% to 19.2% (Figure 4.S2). Receiver operating characteristic (ROC) curves for the prediction models produced by the LR algorithm showed area under the curve values over 0.9 for half of the antibiotics tested (Figure 4.S3a). Though the rules-based approach was biased to produce major errors rather than very major errors in the case of ambiguity, the logistic regression algorithm was not biased towards either error type. Despite this, the very major error rates were similar between the two algorithms (RB: 4.8%, 95% confidence interval (CI) 2.8% to 6.7%; LR: 6.6%,

95% CI 4.0% to 9.2%). Both algorithms correctly identified susceptible and resistant organisms for each of the 12 antibiotics tested.

We next sought to determine if species-specific factors contributed predictive power to the LR algorithm, or if most of the predictive power came from species-independent variables, such as the presence of horizontally transferred resistance genes or mutations in conserved enzymes. To make this distinction we repeated the LR predictions either excluding species as a variable (Figure 4.2) or using species as the only variable (Figure 4.S4). While species information appeared to have some predictive value as a variable on its own (70.7% accurate predictions, with 18.4% major errors and 10.9% very major errors), it was not independent from the predictive value of the factors shared between species, as the accuracy percentage of the complete model dropped only a small amount when the species variable was excluded (90.3% with a major error rate of 3.3% and a very major error rate of 6.4%, Figure 4.S3b).

Alternative input variables for Ciprofloxacin: The RB algorithm for predicting ciprofloxacin resistance was based on the quinolone resistance determining region (QRDR) of the *gyrA* and *gyrB* DNA gyrase and *parC* topoisomerase genes, but several other genes are known to have an effect on fluoroquinolone resistance levels. These include the *qnr* protection protein and the *oqx* efflux pumps, both of which were present in >29% of the genomes tested in this study. Unlike resistant gyrase or topoisomerase genes, the effects of these resistance factors are transferrable between strains, but these genes do not seem to provide enough resistance to surpass clinical breakpoints on their own (31). Our experiments support that conclusion as inclusion of these genes as resistance inputs for the RB algorithm consistently lowered accuracy (Figure 4.3a). To determine if a particular subset of mutations could be predictive of most of the resistance, we performed attribute selection on the inputs to the LR algorithm. Attribute selection identified

four mutations to be predictive, but using only these mutations as inputs greatly reduced the predictive power of the LR algorithm (Figure 4.3b).

Functional identification of resistance genes: To determine if novel antibiotic resistance genes, not identified by Resfams, contributed to our very major error rate, we used functional metagenomics to identify the transferrable resistance genes present in the isolates. We sequenced functionally-selected genomic DNA fragments conferring resistance using short read sequencing and aligned the reads to each of the 78 genomes individually (Figure 4.4). No resistance genes to ciprofloxacin were identified by functional metagenomics. The only antibiotic class for which we identified resistance genes by functional metagenomics that were not identified from the genome sequence were the beta-lactams. The majority of genes in this category were *ramA* transcriptional regulators which are known to confer resistance when overexpressed in *E. coli* (115), but which are a part of the baseline resistance levels of their natural *K. pneumoniae* and *Enterobacter* sp. hosts. Similarly, the chromosomal *ampC* carried by *E. coli* were ignored from the genome sequence because they are known to be repressed, but when overexpressed (as in functional metagenomics) they provide broad-spectrum resistance. Accordingly, the functional metagenomic results supported the whole genome resistance gene identification, and did not identify any novel genes that could have caused errors in the predictions.

Doxycycline repeat analysis: Doxycycline was unique among the antibiotics tested in that all of the errors for each prediction algorithm were very major errors (Figure 4.2). Since no novel tetracycline resistance genes were detected in the functional metagenomic experiment (Figure 4.4), we performed repeat phenotypic AST on the erroneous isolates with three independent testers (Table 4.2). All but two of these isolates were re-classified as susceptible by all three

testers, one was classified as susceptible by two of the three testers, and the last retained its resistant classification for all three testers.

Quantitative predictions of resistance: One advantage of using machine learning rather than empirical rules is in making predictions about an outcome for which rules have not been defined. Specifically, exact values for the minimum inhibitory concentration (MIC) of a pathogen against an antibiotic could be predicted from its resistance gene content, allowing clinicians to assess the degree of resistance, as they can with *in vitro* phenotype-based diagnostics. We used the size of the zones of inhibition (ZOI) as the output from a random forest (RF) machine learning algorithm. Gentamicin ZOI were best predicted by this method, with a correlation of determination (R^2) value of 0.86, but the R^2 values for seven of the eleven other antibiotics were below 0.7 (Figure 4.S5).

4.5 Discussion

For integration into the routine clinical laboratory setting, genome-based susceptibility predictions for bacterial infections will require a fully automated interpretation pipeline. While it could be initially less complex to create a system where a human translates the resistance gene list from WGS into a predicted phenotype, a computer can perform this task more quickly, consistently, objectively, and quantitatively. The question is whether to use an algorithm based approach populated with current scientific knowledge, such as our RB algorithm, or one that uses machine learning, such as our LR or RF approaches. Both methods performed similarly in this study, but there are three major reasons to believe that machine learning will yield a more viable long-term approach. First, the *Enterobacteriaceae* represent a best-case pathogen family for use of the RB algorithm, since it is the bacterial family about which we have the most knowledge of molecular and genomic characterization of antibiotic resistance. WGS-based AST would be

useful for many other pathogens with very different resistance patterns particularly *Mycobacterium tuberculosis* infections, as has been suggested previously (116, 117). The LR and RF algorithms, on the other hand, should perform equally well on any bacterial pathogen for which we know the types of resistance factors and have sufficient numbers of characterized, banked isolates for training. Second, as our knowledge of resistance increases or as new resistance genes enter the pathogenic population, the RB algorithm will need to have new rules added, and it will become increasingly complex. To update the LR algorithm though, will only require adding new genes to the input list, and the minor increases in complexity caused by additional inputs will balanced by the increased size of the training set that will come from sequencing and phenotyping more isolates. Third, machine learning algorithms can be used to predict quantitative measures of resistance, such as MIC or ZOI, providing clinically useful information not currently available from many rapid AST methods. Properly training a machine learning algorithm to predict MIC or ZOI would require a larger training set than was possible in the scope of this study, but our preliminary results in this field show promise for this technology.

Any clinically viable method of WGS-based AST will require knowledge of the possible resistance genes present in a pathogen population. Functional metagenomics is a useful technique to inform the presence of resistance genes as it can be used to assay a large number of isolates in a single experiment. Our functional metagenomic assay show that the resistance genes identified from genome sequence are effective within the *Enterobacteriaceae* regardless of specific host species, supporting the minimal effect of species seen in the LR model. This experiment additionally demonstrated that the majority of functional resistance genes in these isolates were also identified from the genome sequence. One of the advantages of functional metagenomics is that it can be used to identify the spectrum of resistance within an antibiotic

class. For instance, in the aminoglycosides, only genes that provide resistance to gentamicin were identified by functional metagenomics, though the full set of aminoglycoside resistance genes could be seen in the genome sequence. By performing selections on multiple selections in the same class, researchers can use functional metagenomics to define the spectrum of a novel resistance gene as it is identified. One downside of functional metagenomics is that the stochastic nature of shotgun cloning can cause genes to be missed, as was the case for a handful of genes in the beta-lactam, chloramphenicol, and tetracycline selections (Figure 4.4). In the case of trimethoprim, the resistance genes identified from both the functional selections and genome sequence were the target of trimethoprim, dihydrofolate reductase (DHFR), a subset of which are naturally resistant to trimethoprim. We anticipated that any DHFR, whether naturally susceptible or resistant, could provide resistance when heterologously overexpressed in a functional metagenomic selection. We were therefore surprised to note that a much lower proportion of *E. coli* DHFR genes were identified in the functional selections than DHFR genes from the other three species.

Despite the high accuracy of our predictions, they did not achieve the standards required to meet the rigor required for FDA clearance standards for an antibiotic resistance *in vitro* diagnostic assay. Given the potential benefits of faster diagnostic techniques, and the several clinical goals for which genome sequencing may be used, taking steps to improve the predictive accuracy achievable of WGS-based AST demonstrated in this study would have great clinical benefit. Some of these improvements would need to be technical. Draft genome sequences are currently more cost efficient to produce than complete genomes, but they may have assembly breaks within key genes, leading to prediction errors. The impact of this limitation was most obvious in our ciprofloxacin predictions, where incompletely assembled QRDR regions resulted

in predictions that were not determined (Figure 4.2). This type of error was more difficult to identify when the incompletely assembled gene was part of the accessory genome but we anticipate that some of our very major errors in each antibiotic condition resulted from incomplete genome assembly. A second technical challenge is the choice of gold standard technique by which to evaluate and refine predictive algorithms. The accepted, inherent error in any *in vitro* susceptibility testing method is a variance of plus or minus one doubling dilution for the minimum inhibitory concentration value. While we cannot evaluate the complete effect that the variability inherent in phenotype-based AST has on our analysis, our repeat analysis of the doxycycline testing showed that some of the disagreement between our predictions and the *in vitro* susceptibility testing were due to variable interpretations of results for isolates near the border between susceptible and non-susceptible. This may have been particularly true for the doxycycline predictions because of the large number of isolates near that border, but we hypothesize that it may have had some effect on all of our predictions.

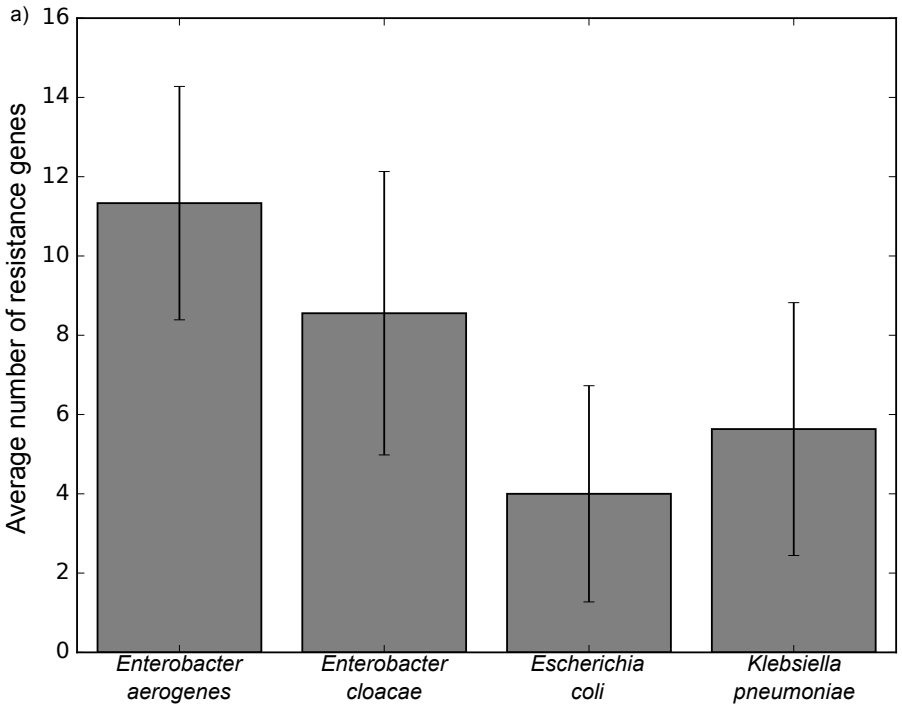
There were also some antibiotic specific errors that resulted from our still incomplete knowledge of antibiotic resistance in the *Enterobacteriaeceae*. To trace these errors we focused on the RB algorithm because the processes by which it makes a prediction are explicitly defined. For the beta-lactams, the majority of the major errors were caused by the RB algorithm's conservative assumptions regarding the resistance spectrums of CTX-M-15 and novel TEM variants. Specifically, it has been reported that CTX-M-15 has activity against ceftazidime and variable activity against cefepime (118), and the RB algorithm predicted resistance against these two antibiotics when CTX-M-15 was present; however, in several cases, isolates encoding CTX-M-15 were susceptible to ceftazidime or cefepime, leading to major errors. The TEM beta-lactamase family is highly diverse, and includes several variants that give resistance to

ceftazidime and third generation cephalosporins such as ceftriaxone (118). Since we could not predict the spectrum of a novel TEM variant in advance, the RB algorithm assumed the highest spectrum (i.e. most conservative prediction) known for a natural TEM variant, though in this set we did not observe any cases where the TEM variant appeared to be granting resistance to later generation cephalosporins. In contrast the very major errors across the beta-lactams were largely caused by outer membrane porin deletions extending the spectrum of beta-lactamases encoded by the pathogen. We could identify these porin deletions by comparing detected porin genes between isolates previously determined to be closely related (80), but we could not identify resistance causing porin deletions without that phylogenetic information. Cefotetan represented a special case as only four of the ten very major errors in predicting susceptibility were caused by porin deletions, while the cause of its remaining very major errors could not be determined. The major errors in the ciprofloxacin predictions also had errors that require further study to resolve, since they resulted from susceptible isolates carrying the canonical *gyrA* QRDR mutation S83L (119).

The pathogenic isolates analyzed in this study were enriched for antibiotic resistance, both a strength and a weakness of this study set. Carbapenem resistant *Enterobacteriaceae*, which represented 24 of our 78 isolates, have been designated one of the most urgent antibiotic resistant threats in the U.S. today (1), and so our set represents some of the most challenging pathogens clinicians are faced with from a treatment perspective. The high degree of accuracy of our predictions with this set shows that WGS-based AST will remain a viable option for rapid diagnostics even as the prevalence of antibiotic resistance continues to increase. At the same time, predictions models built by the LR algorithm on this set could potentially overestimate resistance in average clinical isolates, and they should be rebuilt on a more representative set.

The definition of a representative set will be highly variable, based on geography, patient population and specimen type. Another important strength of this study was our comparison of an algorithm based on prior knowledge with algorithms based on machine learning. Though the machine learning techniques performed better and showed potential for extension to quantitative predictions, the RB algorithm allowed us to identify specific sources of error in its predictions. Reducing the errors for which we can identify the source in this way will go a long way towards making WGS-based AST clinically viable. One limitation of WGS-based AST, which we were not able to account for, was differences in expression level of resistance genes in different contexts. This is known to be important for the spectrum and resistance level of beta-lactamases (120, 121), and it likely affected our quantitative predictions of resistance. Further studies will be necessary to determine the extent to which expression level can be accounted for in WGS-based AST. The fastest way to improve our predictive power, especially for machine learning approaches, will be to publish more genomes with matched phenotypic AST data. As the available training set increases, the performance of WGS-based AST will improve.

4.6 Figures


a)
Average number of resistance genes
0
2
4
6
8
10
12
14
16
Enterobacter aerogenes
Enterobacter cloacae
Escherichia coli
Klebsiella pneumoniae

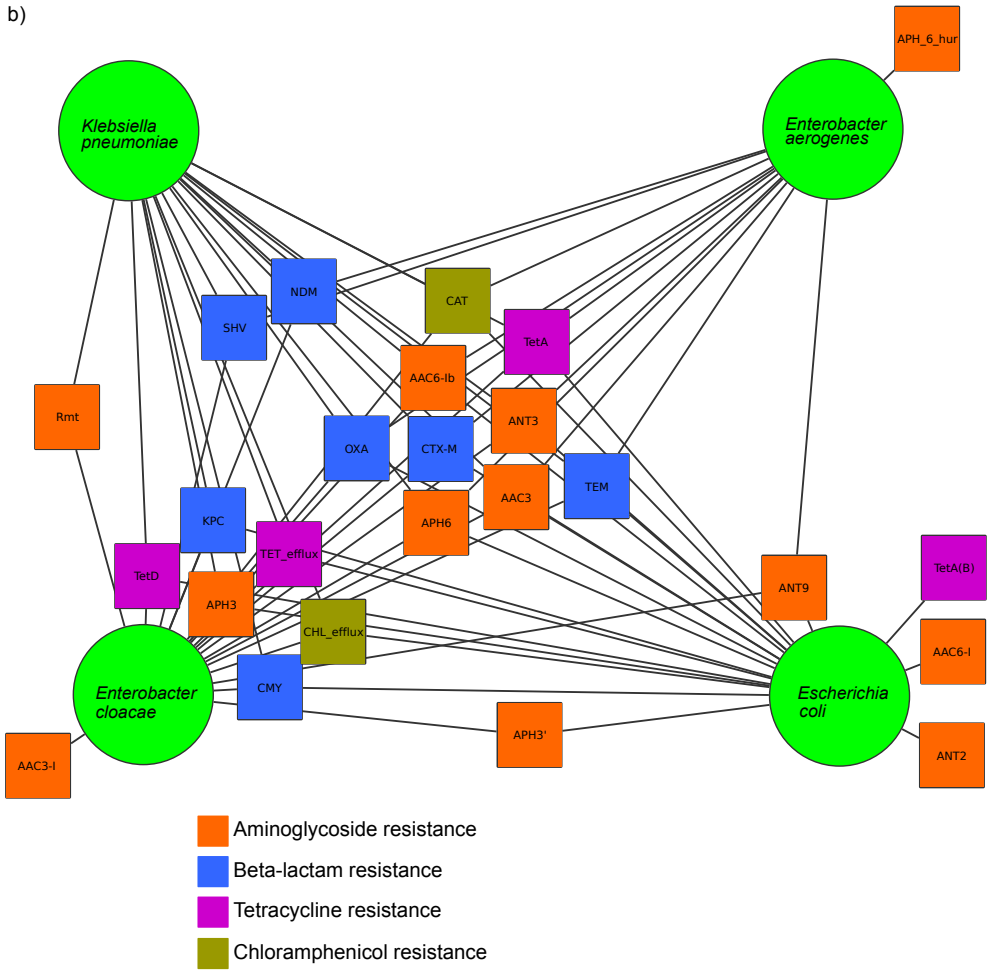

b)
Klebsiella pneumoniae
Enterobacter aerogenes
Enterobacter cloacae
Escherichia coli
APH_6_hur
NDM
SHV
CAT
TetA
AAC6-Ib
Rmt
ANT3
OXA
CTX-M
TEM
AAC3
KPC
APH6
TET_efflux
TetD
APH3
CHL_efflux
ANT9
TetA(B)
AAC6-I
CMY
APH3'
AAC3-I
ANT2
Aminoglycoside resistance
Beta-lactam resistance
Tetracycline resistance
Chloramphenicol resistance

Figure 4.1

Resistance gene content and sharing of the tested isolates. a) Average number of resistance genes per genome in each of the four species tested. b) Network diagram demonstrating gene sharing between the four species. Each square represents a resistance gene family colored by class of antibiotic. A line between a gene family and a species indicates that the resistance gene family was found within at least one isolate from that species. Gene families were manually clustered based on the species in which they were found.

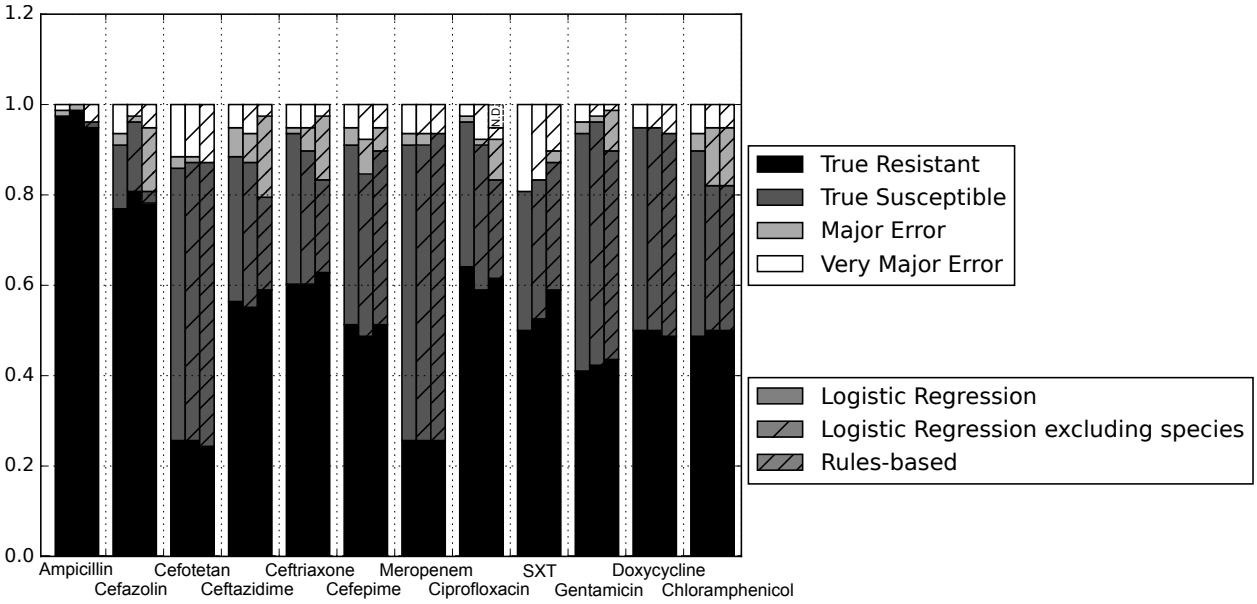


Figure 4.2

Prediction accuracy of WGS-based AST algorithms. True Resistant: both the prediction algorithm and the gold standard AST returned “resistant”. True Susceptible: both the prediction algorithm and gold standard AST returned “susceptible”. Major Error: the prediction algorithm returned “resistant” while the gold standard AST returned “susceptible”. Very Major Error: the prediction algorithm returned “susceptible” while the gold standard AST returned “resistant”.

SXT = trimethoprim-sulfamethoxazole. N.D. Susceptibility could not be predicted for this antibiotic and these isolates.

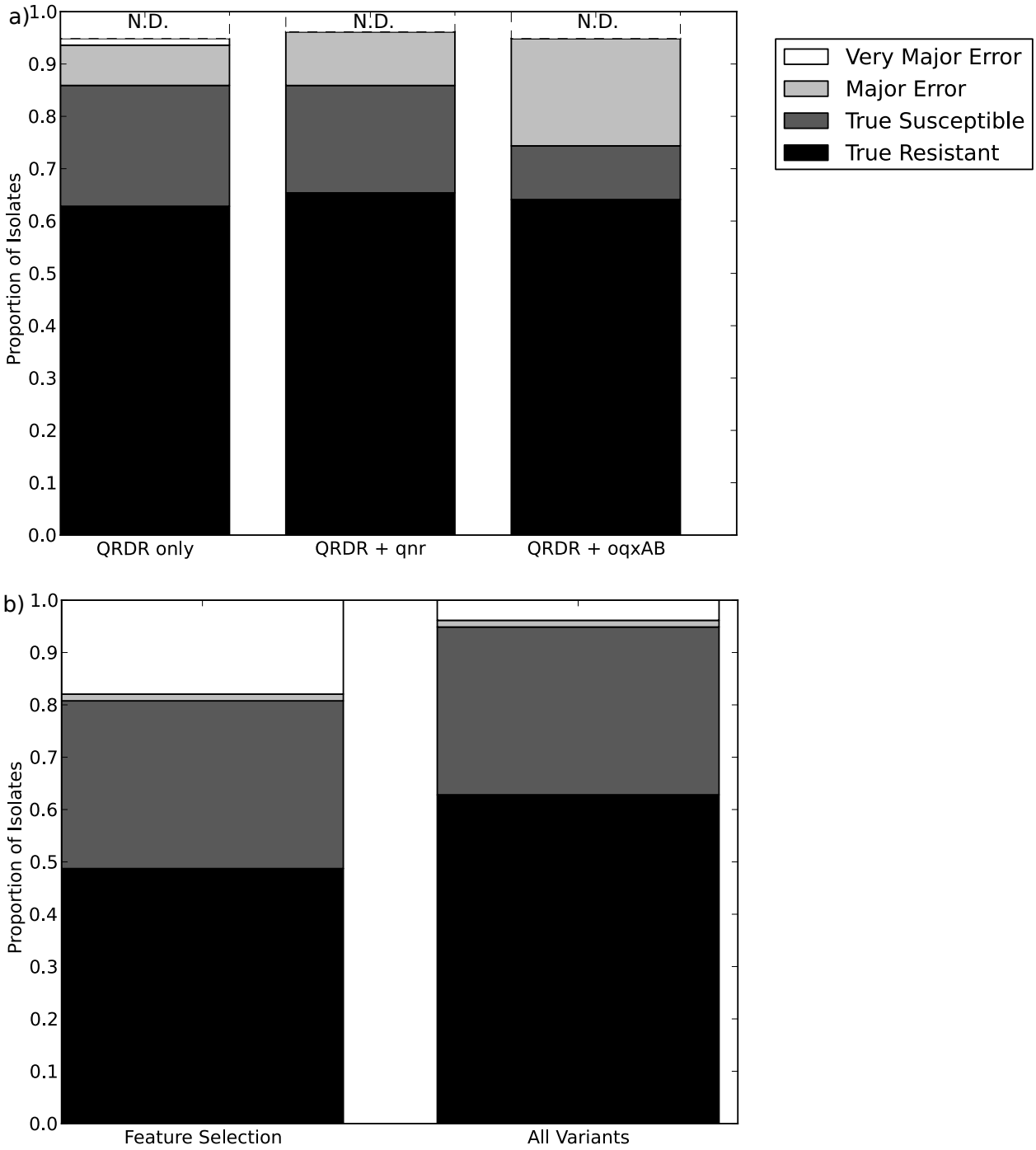

a)
N.D.
N.D.
N.D.
Proportion of Isolates
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
QRDR only
QRDR + qnr
QRDR + oqxAB
Very Major Error
Major Error
True Susceptible
True Resistant
b)
Proportion of Isolates
Feature Selection
All Variants

Figure 4.3

Alternative predictions schemes for ciprofloxacin. a) Rules-based algorithm predictions for ciprofloxacin using QRDR mutations only, or with *qnr* and *oqxA* and *oqxB* as resistance genes. N.D. Susceptibility could not be predicted for this antibiotic and these isolates. b) Logarithmic Regression algorithm predictions using all QRDR mutations, or using only those shown by feature selection to have the greatest impact on the phenotype.

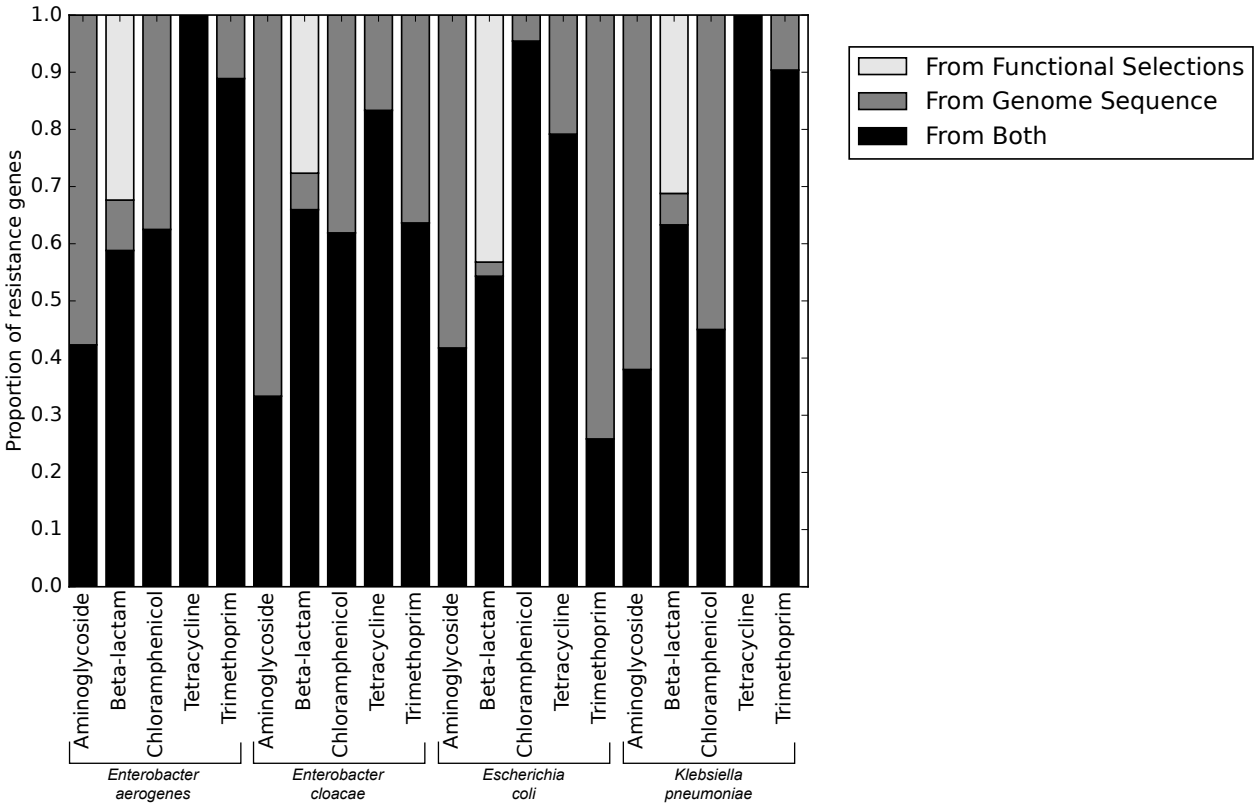


Figure 4.4

Method used to identify antibiotic resistance gene, separated by species.

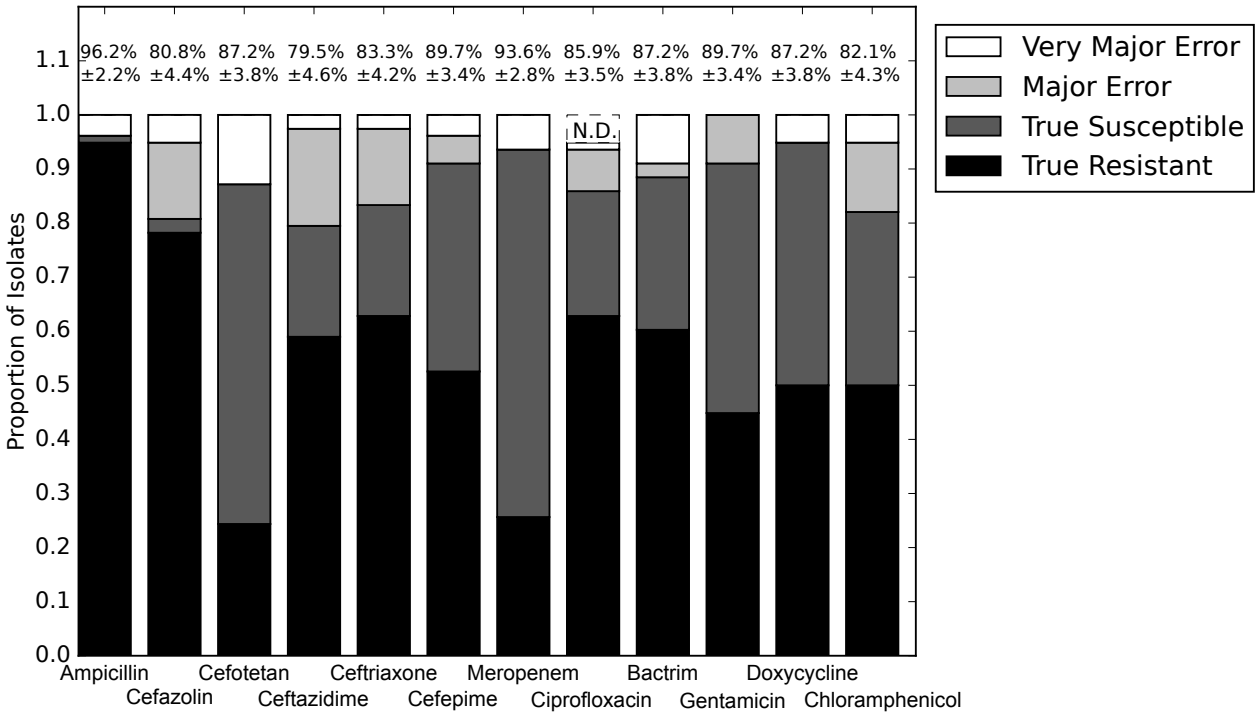


Figure 4.S1

Prediction accuracy for RB algorithm alone. Percentages above bars represent percent accurate predictions and standard error for accuracy percentage. SXT = trimethoprim-sulfamethoxazole. N.D. Susceptibility could not be predict for this antibiotic and these isolate.

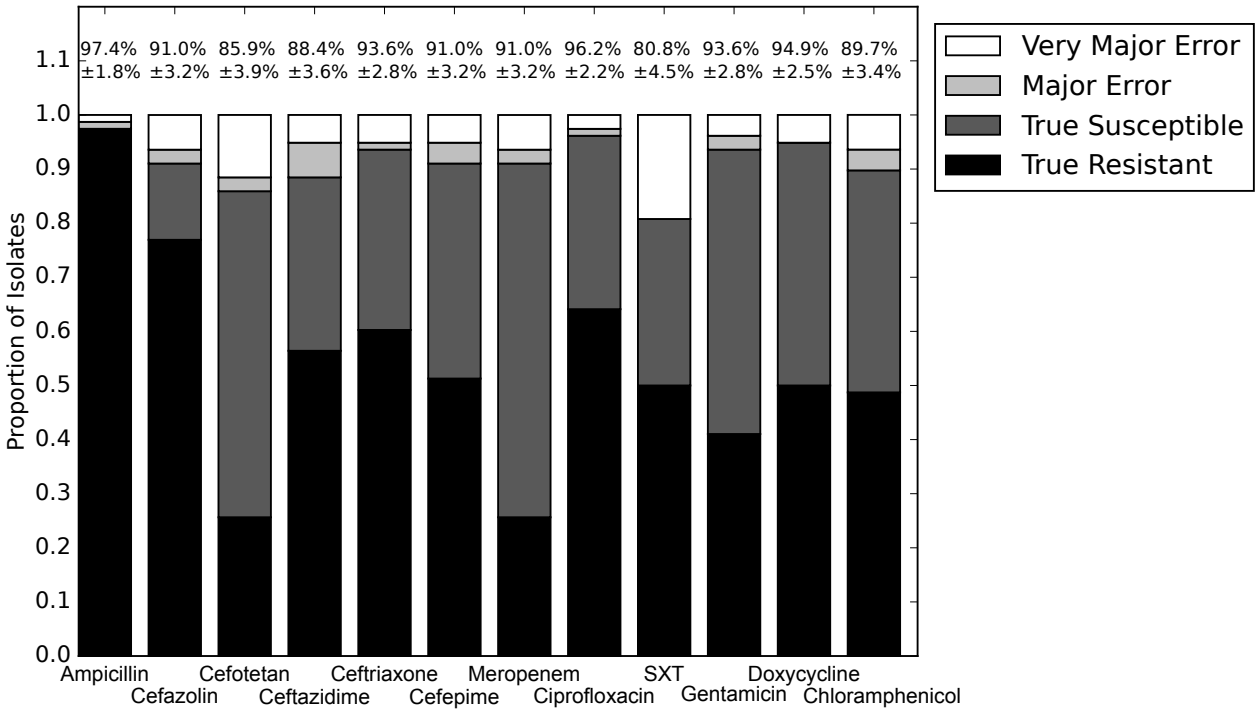
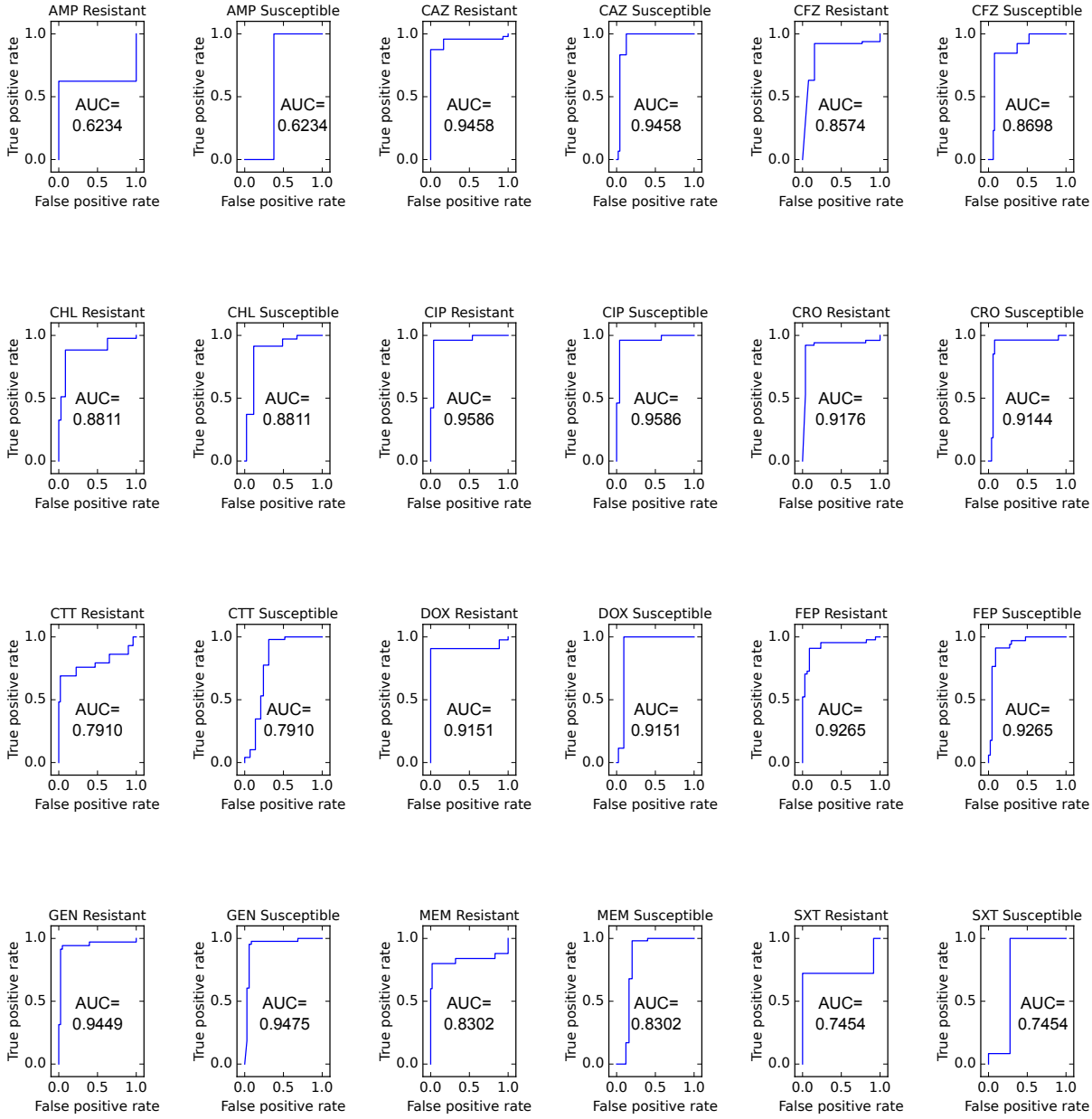


Figure 4.S2

Prediction for the LR algorithm alone. Percentages above bars represent percent accurate predictions and standard error for accuracy percentage. SXT = trimethoprim-sulfamethoxazole.

a)


AMP Resistant
AUC= 0.6234
AMP Susceptible
AUC= 0.6234
CAZ Resistant
AUC= 0.9458
CAZ Susceptible
AUC= 0.9458
CFZ Resistant
AUC= 0.8574
CFZ Susceptible
AUC= 0.8698
CHL Resistant
AUC= 0.8811
CHL Susceptible
AUC= 0.8811
CIP Resistant
AUC= 0.9586
CIP Susceptible
AUC= 0.9586
CRO Resistant
AUC= 0.9176
CRO Susceptible
AUC= 0.9144
CTT Resistant
AUC= 0.7910
CTT Susceptible
AUC= 0.7910
DOX Resistant
AUC= 0.9151
DOX Susceptible
AUC= 0.9151
FEP Resistant
AUC= 0.9265
FEP Susceptible
AUC= 0.9265
GEN Resistant
AUC= 0.9449
GEN Susceptible
AUC= 0.9475
MEM Resistant
AUC= 0.8302
MEM Susceptible
AUC= 0.8302
SXT Resistant
AUC= 0.7454
SXT Susceptible
AUC= 0.7454
True positive rate
False positive rate
0.0
0.5
1.0

b)

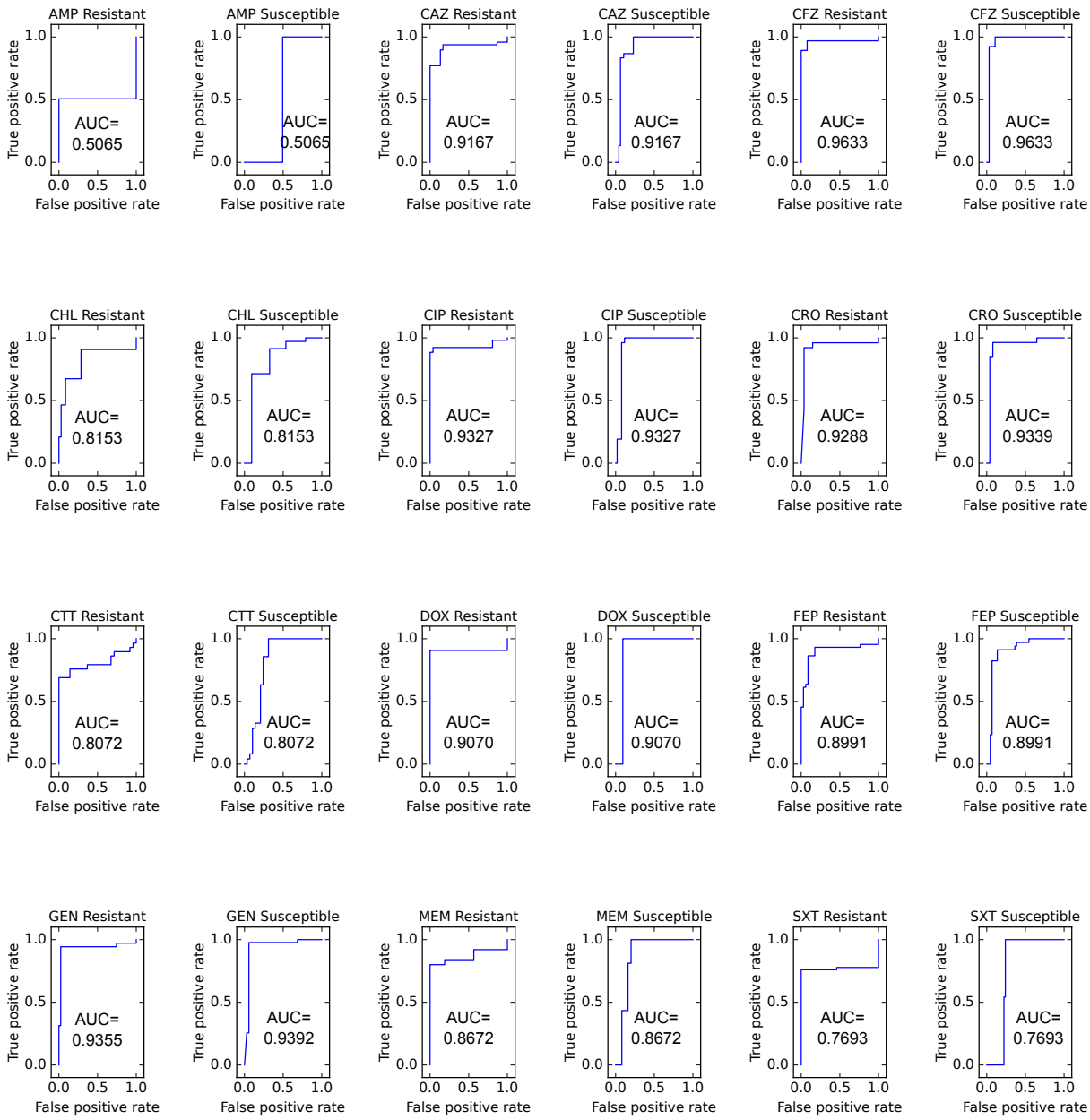


Figure 4.S3

ROC curves for predicting susceptible and resistant isolates for each antibiotic using the LR algorithm a) including and b) not including species as an input. Area under the curve (AUC) is given for each ROC curve.

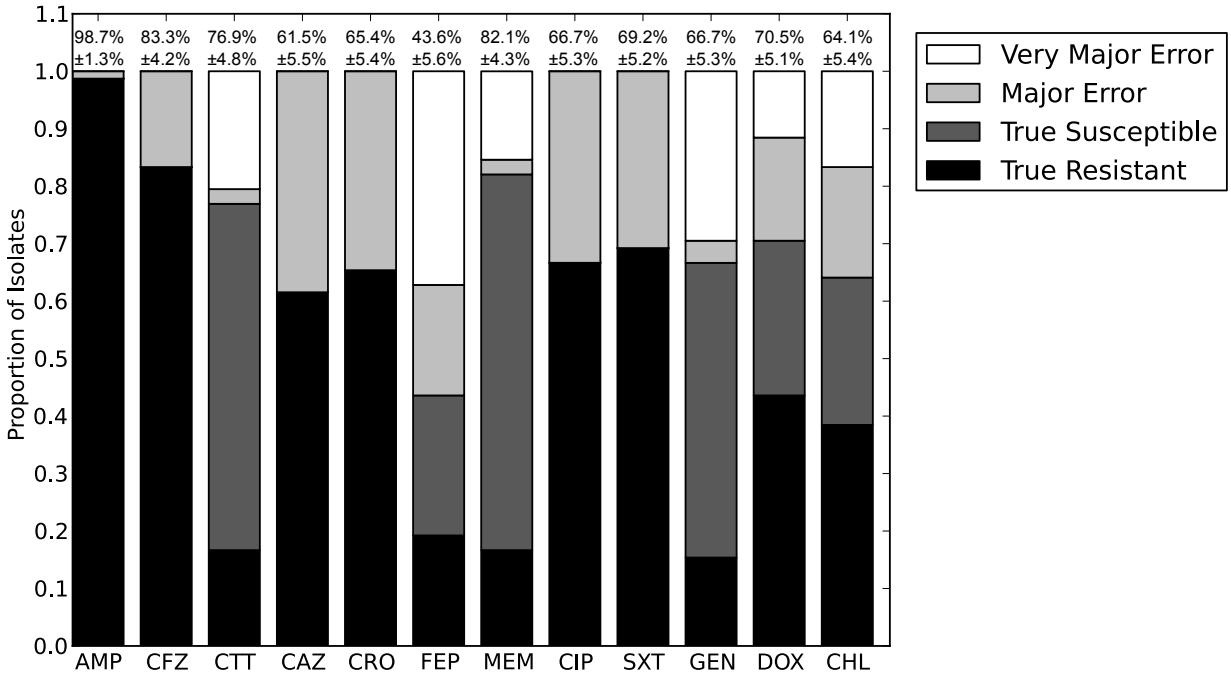


Figure 4.S4

Prediction accuracy for LR algorithm using species as the only input variable. AMP = ampicillin. CFZ = cefazolin. CTT = cefotetan. CRO = ceftriaxone. FEP = cefepime. MEM = meropenem. Cip = ciprofloxacin. SXT = trimethoprim-sulfamethoxazole. GEN = gentamicin. DOX = doxycycline. CHL = chloramphenicol.

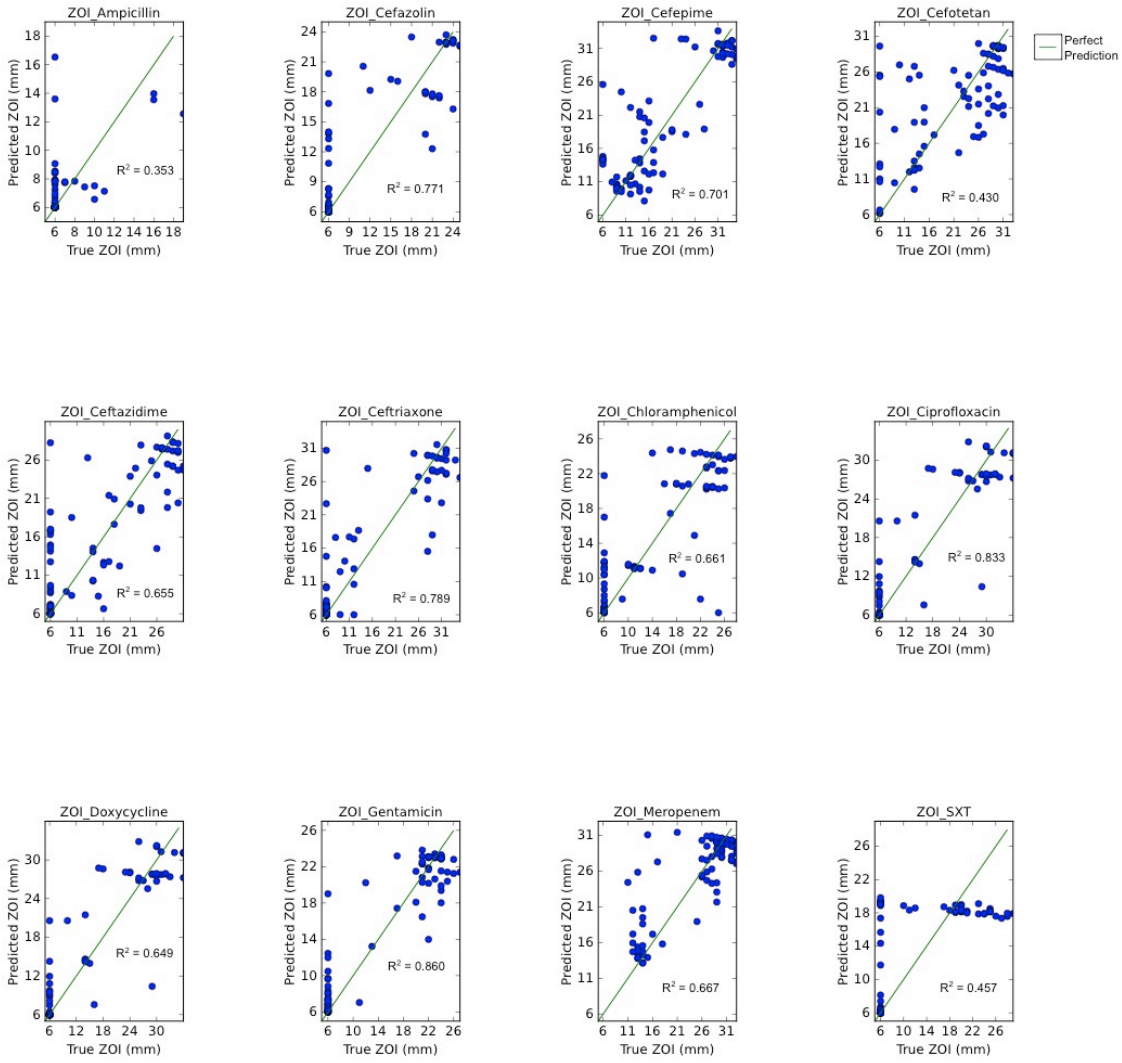


Figure 4.S5

Performance of Random Forest machine learning algorithm in predicting size of zone of inhibition (ZOI) from genomic data, for each of the twelve antibiotics tested. Coefficients of determination and lines indicating perfect prediction are given for each antibiotic.

4.7 Tables

Isolate	Species	AMP ZOI	AM Interpretation	CAZ ZOI	CZ Interpretation
PH100	*Escherichia coli*	6	Resistant	20	Resistant
PH101-2	*Escherichia coli*	6	Resistant	6	Resistant
PH105	*Escherichia coli*	6	Resistant	21	Intermediate
PH108	*Escherichia coli*	6	Resistant	6	Resistant
PH114	*Escherichia coli*	6	Resistant	6	Resistant
PH118	*Escherichia coli*	6	Resistant	6	Resistant
PH129	*Escherichia coli*	6	Resistant	6	Resistant
PH135	*Escherichia coli*	6	Resistant	6	Resistant
PH141	*Escherichia coli*	16	Intermediate	22	Intermediate
PH143	*Escherichia coli*	6	Resistant	6	Resistant
PH151-2	*Escherichia coli*	6	Resistant	6	Resistant
PH156-1	*Escherichia coli*	6	Resistant	6	Resistant
PH18	*Escherichia coli*	6	Resistant	6	Resistant
PH20	*Escherichia coli*	6	Resistant	6	Resistant
PH31	*Escherichia coli*	6	Resistant	6	Resistant
PH39	*Escherichia coli*	6	Resistant	21	Intermediate
PH51	*Escherichia coli*	6	Resistant	15	Resistant
PH5	*Escherichia coli*	6	Resistant	6	Resistant
PH85	*Escherichia coli*	6	Resistant	22	Resistant
PH90	*Escherichia coli*	6	Resistant	6	Resistant
PH92-1	*Escherichia coli*	6	Resistant	20	Intermediate
PH93	*Escherichia coli*	6	Resistant	20	Resistant
PH94	*Escherichia coli*	6	Resistant	6	Resistant
PH98	*Escherichia coli*	6	Resistant	20	Resistant
WU31	*Escherichia coli*	6	Resistant	6	Resistant
WU32	*Escherichia coli*	6	Resistant	6	Resistant
WU33	*Escherichia coli*	6	Resistant	6	Resistant
WU34	*Escherichia coli*	6	Resistant	6	Resistant
WU35	*Escherichia coli*	16	Intermediate	16	Intermediate
WU40	*Escherichia coli*	6	Resistant	11	Resistant
WU43	*Escherichia coli*	19	Susceptible	24	Susceptible
WU44	*Escherichia coli*	6	Resistant	12	Resistant
WU45	*Escherichia coli*	6	Resistant	21	Intermediate

PH112-2	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH113	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH134	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH138-2	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH63	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH84-2	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH112-1	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH125	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH158	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH23	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH24-2	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH82	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU26	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU27	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU29	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH102	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH10	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH11	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH124	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH12	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH139	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH150-2	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH152	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH24-1	*Klebsiella pneumoniae*	6	Resistant	21	Intermediate
PH25	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH28-1	*Klebsiella pneumoniae*	6	Resistant	23	Susceptible
PH38-1	*Klebsiella pneumoniae*	10	Resistant	23	Susceptible
PH40	*Klebsiella pneumoniae*	9	Resistant	23	Susceptible
PH44	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH49-2	*Klebsiella pneumoniae*	8	Resistant	24	Susceptible

PH72	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH73	*Klebsiella pneumoniae*	6	Resistant	24	Susceptible
PH88	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH9	*Klebsiella pneumoniae*	6	Resistant	18	Resistant
WU10	*Klebsiella pneumoniae*	6	Resistant	25	Susceptible
WU12	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU18	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU21	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU23	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU2	*Klebsiella pneumoniae*	11	Resistant	25	Susceptible
WU3	*Klebsiella pneumoniae*	6	Resistant	24	Susceptible
WU6	*Klebsiella pneumoniae*	7	Resistant	22	Susceptible
WU7	*Klebsiella pneumoniae*	6	Resistant	23	Susceptible
WU8	*Klebsiella pneumoniae*	7	Resistant	23	Susceptible
WU9	*Klebsiella pneumoniae*	10	Resistant	24	Susceptible
Isolate	Species	CTT ZOI	CTT Interpretation	CAZ ZOI	CAZ Interpretation
PH100	*Escherichia coli*	30	Susceptible	28	Susceptible
PH101-2	*Escherichia coli*	31	Susceptible	16	Resistant
PH105	*Escherichia coli*	33	Susceptible	30	Susceptible
PH108	*Escherichia coli*	6	Resistant	6	Resistant
PH114	*Escherichia coli*	25	Susceptible	14	Resistant
PH118	*Escherichia coli*	31	Susceptible	17	Resistant
PH129	*Escherichia coli*	26	Susceptible	14	Resistant
PH135	*Escherichia coli*	30	Susceptible	23	Susceptible
PH141	*Escherichia coli*	28	Susceptible	28	Susceptible
PH143	*Escherichia coli*	6	Resistant	6	Resistant
PH151-2	*Escherichia coli*	29	Susceptible	19	Intermediate
PH156-1	*Escherichia coli*	26	Susceptible	16	Resistant
PH18	*Escherichia coli*	27	Susceptible	17	Resistant
PH20	*Escherichia coli*	6	Resistant	6	Resistant
PH31	*Escherichia coli*	28	Susceptible	18	Intermediate
PH39	*Escherichia coli*	32	Susceptible	31	Susceptible
PH51	*Escherichia coli*	28	Susceptible	26	Susceptible
PH5	*Escherichia coli*	28	Susceptible	21	Susceptible
PH85	*Escherichia coli*	30	Susceptible	29	Susceptible
PH90	*Escherichia coli*	29	Susceptible	23	Susceptible

PH92-1	*Escherichia coli*	31	Susceptible	31	Susceptible
PH93	*Escherichia coli*	31	Susceptible	29	Susceptible
PH94	*Escherichia coli*	27	Susceptible	18	Intermediate
PH98	*Escherichia coli*	28	Susceptible	25	Susceptible
WU31	*Escherichia coli*	9	Resistant	10	Resistant
WU32	*Escherichia coli*	14	Intermediate	14	Resistant
WU33	*Escherichia coli*	15	Intermediate	14	Resistant
WU34	*Escherichia coli*	12	Resistant	6	Resistant
WU35	*Escherichia coli*	13	Intermediate	13	Resistant
WU40	*Escherichia coli*	22	Susceptible	21	Susceptible
WU43	*Escherichia coli*	30	Susceptible	30	Susceptible
WU44	*Escherichia coli*	21	Susceptible	22	Susceptible
WU45	*Escherichia coli*	29	Susceptible	28	Susceptible
PH112-2	*Enterobacter aerogenes*	13	Intermediate	6	Resistant
PH113	*Enterobacter aerogenes*	14	Intermediate	6	Resistant
PH134	*Enterobacter aerogenes*	13	Intermediate	6	Resistant
PH138-2	*Enterobacter aerogenes*	13	Intermediate	6	Resistant
PH63	*Enterobacter aerogenes*	13	Intermediate	6	Resistant
PH84-2	*Enterobacter aerogenes*	12	Resistant	6	Resistant
PH112-1	*Enterobacter cloacae* complex	26	Susceptible	16	Resistant
PH125	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH158	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH23	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH24-2	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH82	*Enterobacter cloacae* complex	22	Susceptible	6	Resistant
WU26	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU27	*Enterobacter cloacae* complex	13	Resistant	10	Resistant
WU29	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH102	*Klebsiella pneumoniae*	10	Resistant	6	Resistant
PH10	*Klebsiella pneumoniae*	23	Susceptible	6	Resistant
PH11	*Klebsiella pneumoniae*	6	Resistant	6	Resistant

PH124	*Klebsiella pneumoniae*	27	Susceptible	15	Resistant
PH12	*Klebsiella pneumoniae*	9	Resistant	6	Resistant
PH139	*Klebsiella pneumoniae*	26	Susceptible	6	Resistant
PH150-2	*Klebsiella pneumoniae*	23	Susceptible	6	Resistant
PH152	*Klebsiella pneumoniae*	24	Susceptible	6	Resistant
PH24-1	*Klebsiella pneumoniae*	28	Susceptible	26	Susceptible
PH25	*Klebsiella pneumoniae*	24	Susceptible	6	Resistant
PH28-1	*Klebsiella pneumoniae*	29	Susceptible	27	Susceptible
PH38-1	*Klebsiella pneumoniae*	29	Susceptible	26	Susceptible
PH40	*Klebsiella pneumoniae*	30	Susceptible	27	Susceptible
PH44	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH49-2	*Klebsiella pneumoniae*	30	Susceptible	28	Susceptible
PH72	*Klebsiella pneumoniae*	17	Susceptible	9	Resistant
PH73	*Klebsiella pneumoniae*	30	Susceptible	28	Susceptible
PH88	*Klebsiella pneumoniae*	24	Susceptible	6	Resistant
PH9	*Klebsiella pneumoniae*	26	Susceptible	23	Susceptible
WU10	*Klebsiella pneumoniae*	31	Susceptible	30	Susceptible
WU12	*Klebsiella pneumoniae*	15	Intermediate	6	Resistant
WU18	*Klebsiella pneumoniae*	15	Intermediate	6	Resistant
WU21	*Klebsiella pneumoniae*	14	Intermediate	6	Resistant
WU23	*Klebsiella pneumoniae*	13	Intermediate	6	Resistant
WU2	*Klebsiella pneumoniae*	31	Susceptible	30	Susceptible
WU3	*Klebsiella pneumoniae*	30	Susceptible	29	Susceptible
WU6	*Klebsiella pneumoniae*	31	Susceptible	29	Susceptible
WU7	*Klebsiella pneumoniae*	29	Susceptible	27	Susceptible
WU8	*Klebsiella pneumoniae*	30	Susceptible	29	Susceptible
WU9	*Klebsiella pneumoniae*	30	Susceptible	30	Susceptible
Isolate	Species	CRO ZOI	CRO Interpretation	FEP ZOI	FEP Interpretation
PH100	*Escherichia coli*	29	Susceptible	27	Susceptible
PH101-2	*Escherichia coli*	6	Resistant	16	Intermediate
PH105	*Escherichia coli*	35	Susceptible	35	Susceptible
PH108	*Escherichia coli*	6	Resistant	17	Intermediate
PH114	*Escherichia coli*	6	Resistant	14	Resistant
PH118	*Escherichia coli*	6	Resistant	19	Susceptible
PH129	*Escherichia coli*	6	Resistant	13	Resistant
PH135	*Escherichia coli*	6	Resistant	21	Susceptible
PH141	*Escherichia coli*	31	Susceptible	32	Susceptible

PH143	*Escherichia coli*	6	Resistant	6	Resistant
PH151-2	*Escherichia coli*	6	Resistant	17	Intermediate
PH156-1	*Escherichia coli*	9	Resistant	15	Intermediate
PH18	*Escherichia coli*	6	Resistant	14	Resistant
PH20	*Escherichia coli*	6	Resistant	6	Resistant
PH31	*Escherichia coli*	6	Resistant	16	Intermediate
PH39	*Escherichia coli*	32	Susceptible	35	Susceptible
PH51	*Escherichia coli*	28	Susceptible	30	Susceptible
PH5	*Escherichia coli*	12	Resistant	21	Susceptible
PH85	*Escherichia coli*	32	Susceptible	34	Susceptible
PH90	*Escherichia coli*	6	Resistant	24	Susceptible
PH92-1	*Escherichia coli*	31	Susceptible	32	Susceptible
PH93	*Escherichia coli*	30	Susceptible	32	Susceptible
PH94	*Escherichia coli*	6	Resistant	19	Susceptible
PH98	*Escherichia coli*	29	Susceptible	26	Susceptible
WU31	*Escherichia coli*	8	Resistant	12	Resistant
WU32	*Escherichia coli*	12	Resistant	17	Resistant
WU33	*Escherichia coli*	10	Resistant	15	Resistant
WU34	*Escherichia coli*	6	Resistant	6	Resistant
WU35	*Escherichia coli*	15	Resistant	23	Susceptible
WU40	*Escherichia coli*	25	Susceptible	32	Susceptible
WU43	*Escherichia coli*	28	Susceptible	34	Susceptible
WU44	*Escherichia coli*	26	Susceptible	31	Susceptible
WU45	*Escherichia coli*	29	Susceptible	33	Susceptible
PH112-2	*Enterobacter aerogenes*	6	Resistant	14	Resistant
PH113	*Enterobacter aerogenes*	6	Resistant	14	Resistant
PH134	*Enterobacter aerogenes*	6	Resistant	12	Resistant
PH138-2	*Enterobacter aerogenes*	6	Resistant	12	Resistant
PH63	*Enterobacter aerogenes*	6	Resistant	12	Resistant
PH84-2	*Enterobacter aerogenes*	6	Resistant	12	Resistant
PH112-1	*Enterobacter cloacae* complex	6	Resistant	16	Intermediate
PH125	*Enterobacter cloacae* complex	6	Resistant	9	Resistant
PH158	*Enterobacter cloacae* complex	6	Resistant	12	Resistant
PH23	*Enterobacter cloacae* complex	6	Resistant	10	Resistant

PH24-2	*Enterobacter cloacae* complex	6	Resistant	9	Resistant
PH82	*Enterobacter cloacae* complex	6	Resistant	9	Resistant
WU26	*Enterobacter cloacae* complex	9	Resistant	13	Resistant
WU27	*Enterobacter cloacae* complex	12	Resistant	14	Resistant
WU29	*Enterobacter cloacae* complex	13	Resistant	10	Resistant
PH102	*Klebsiella pneumoniae*	6	Resistant	15	Intermediate
PH10	*Klebsiella pneumoniae*	6	Resistant	11	Susceptible
PH11	*Klebsiella pneumoniae*	6	Resistant	9	Resistant
PH124	*Klebsiella pneumoniae*	6	Resistant	17	Intermediate
PH12	*Klebsiella pneumoniae*	6	Resistant	9	Resistant
PH139	*Klebsiella pneumoniae*	6	Resistant	15	Intermediate
PH150-2	*Klebsiella pneumoniae*	6	Resistant	10	Resistant
PH152	*Klebsiella pneumoniae*	6	Resistant	8	Resistant
PH24-1	*Klebsiella pneumoniae*	28	Susceptible	28	Susceptible
PH25	*Klebsiella pneumoniae*	6	Resistant	12	Resistant
PH28-1	*Klebsiella pneumoniae*	28	Susceptible	31	Susceptible
PH38-1	*Klebsiella pneumoniae*	28	Susceptible	31	Susceptible
PH40	*Klebsiella pneumoniae*	29	Susceptible	32	Susceptible
PH44	*Klebsiella pneumoniae*	6	Resistant	9	Resistant
PH49-2	*Klebsiella pneumoniae*	30	Susceptible	32	Susceptible
PH72	*Klebsiella pneumoniae*	6	Resistant	14	Resistant
PH73	*Klebsiella pneumoniae*	30	Susceptible	31	Susceptible
PH88	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH9	*Klebsiella pneumoniae*	25	Susceptible	24	Susceptible
WU10	*Klebsiella pneumoniae*	32	Susceptible	35	Susceptible
WU12	*Klebsiella pneumoniae*	12	Resistant	16	Resistant
WU18	*Klebsiella pneumoniae*	11	Resistant	15	Resistant
WU21	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU23	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU2	*Klebsiella pneumoniae*	34	Susceptible	33	Susceptible
WU3	*Klebsiella pneumoniae*	32	Susceptible	33	Susceptible
WU6	*Klebsiella pneumoniae*	31	Susceptible	34	Susceptible
WU7	*Klebsiella pneumoniae*	30	Susceptible	32	Susceptible
WU8	*Klebsiella pneumoniae*	32	Susceptible	34	Susceptible
WU9	*Klebsiella pneumoniae*	32	Susceptible	34	Susceptible

Isolate	Species	CIP ZOI	CIP Interpretation	SXT ZOI	SXT Interpretation
PH100	*Escherichia coli*	30	Susceptible	6	Resistant
PH101-2	*Escherichia coli*	6	Resistant	6	Resistant
PH105	*Escherichia coli*	28	Susceptible	6	Resistant
PH108	*Escherichia coli*	6	Resistant	19	Susceptible
PH114	*Escherichia coli*	6	Resistant	6	Resistant
PH118	*Escherichia coli*	6	Resistant	17	Susceptible
PH129	*Escherichia coli*	6	Resistant	6	Resistant
PH135	*Escherichia coli*	26	Susceptible	29	Susceptible
PH141	*Escherichia coli*	6	Resistant	28	Susceptible
PH143	*Escherichia coli*	6	Resistant	6	Resistant
PH151-2	*Escherichia coli*	6	Resistant	6	Resistant
PH156-1	*Escherichia coli*	6	Resistant	6	Resistant
PH18	*Escherichia coli*	6	Resistant	20	Susceptible
PH20	*Escherichia coli*	6	Resistant	6	Resistant
PH31	*Escherichia coli*	6	Resistant	6	Resistant
PH39	*Escherichia coli*	26	Susceptible	6	Resistant
PH51	*Escherichia coli*	6	Resistant	25	Susceptible
PH5	*Escherichia coli*	6	Resistant	6	Resistant
PH85	*Escherichia coli*	26	Susceptible	6	Resistant
PH90	*Escherichia coli*	6	Resistant	6	Resistant
PH92-1	*Escherichia coli*	30	Susceptible	6	Resistant
PH93	*Escherichia coli*	27	Susceptible	6	Resistant
PH94	*Escherichia coli*	6	Resistant	6	Resistant
PH98	*Escherichia coli*	30	Susceptible	6	Resistant
WU31	*Escherichia coli*	6	Resistant	6	Resistant
WU32	*Escherichia coli*	6	Resistant	6	Resistant
WU33	*Escherichia coli*	36	Susceptible	20	Susceptible
WU34	*Escherichia coli*	6	Resistant	6	Resistant
WU35	*Escherichia coli*	31	Susceptible	6	Resistant
WU40	*Escherichia coli*	34	Susceptible	25	Susceptible
WU43	*Escherichia coli*	36	Susceptible	28	Susceptible
WU44	*Escherichia coli*	6	Resistant	25	Susceptible
WU45	*Escherichia coli*	6	Resistant	19	Susceptible
PH112-2	*Enterobacter aerogenes*	14	Resistant	6	Resistant
PH113	*Enterobacter aerogenes*	14	Resistant	6	Resistant

PH134	*Enterobacter aerogenes*	15	Resistant	6	Resistant
PH138-2	*Enterobacter aerogenes*	14	Resistant	6	Resistant
PH63	*Enterobacter aerogenes*	14	Resistant	6	Resistant
PH84-2	*Enterobacter aerogenes*	14	Resistant	6	Resistant
PH112-1	*Enterobacter cloacae* complex	16	Intermediate	6	Resistant
PH125	*Enterobacter cloacae* complex	6	Resistant	23	Susceptible
PH158	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH23	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH24-2	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
PH82	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU26	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU27	*Enterobacter cloacae* complex	6	Resistant	6	Resistant
WU29	*Enterobacter cloacae* complex	29	Susceptible	24	Susceptible
PH102	*Klebsiella pneumoniae*	18	Intermediate	6	Resistant
PH10	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH11	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH124	*Klebsiella pneumoniae*	10	Resistant	6	Resistant
PH12	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH139	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH150-2	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH152	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH24-1	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH25	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH28-1	*Klebsiella pneumoniae*	30	Susceptible	20	Susceptible
PH38-1	*Klebsiella pneumoniae*	30	Susceptible	20	Susceptible
PH40	*Klebsiella pneumoniae*	30	Susceptible	21	Susceptible
PH44	*Klebsiella pneumoniae*	6	Resistant	11	Intermediate
PH49-2	*Klebsiella pneumoniae*	29	Susceptible	21	Susceptible
PH72	*Klebsiella pneumoniae*	17	Intermediate	6	Resistant
PH73	*Klebsiella pneumoniae*	29	Susceptible	25	Resistant
PH88	*Klebsiella pneumoniae*	6	Resistant	20	Susceptible

PH9	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU10	*Klebsiella pneumoniae*	32	Susceptible	26	Susceptible
WU12	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU18	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
WU21	*Klebsiella pneumoniae*	6	Resistant	18	Susceptible
WU23	*Klebsiella pneumoniae*	23	Susceptible	6	Resistant
WU2	*Klebsiella pneumoniae*	36	Susceptible	19	Susceptible
WU3	*Klebsiella pneumoniae*	32	Susceptible	24	Susceptible
WU6	*Klebsiella pneumoniae*	33	Susceptible	27	Susceptible
WU7	*Klebsiella pneumoniae*	24	Susceptible	12	Intermediate
WU8	*Klebsiella pneumoniae*	24	Susceptible	10	Resistant
WU9	*Klebsiella pneumoniae*	31	Susceptible	23	Susceptible
Isolate	Species	GEN ZOI	GEN Interpretation	DOX ZOI	DOX Interpretation
PH100	*Escherichia coli*	24	Susceptible	6	Resistant
PH101-2	*Escherichia coli*	24	Susceptible	6	Resistant
PH105	*Escherichia coli*	6	Resistant	6	Resistant
PH108	*Escherichia coli*	24	Susceptible	6	Resistant
PH114	*Escherichia coli*	6	Resistant	6	Resistant
PH118	*Escherichia coli*	6	Resistant	6	Resistant
PH129	*Escherichia coli*	6	Resistant	10	Resistant
PH135	*Escherichia coli*	25	Susceptible	20	Susceptible
PH141	*Escherichia coli*	26	Susceptible	20	Susceptible
PH143	*Escherichia coli*	6	Resistant	6	Resistant
PH151-2	*Escherichia coli*	6	Resistant	13	Intermediate
PH156-1	*Escherichia coli*	20	Susceptible	19	Susceptible
PH18	*Escherichia coli*	21	Susceptible	11	Intermediate
PH20	*Escherichia coli*	6	Resistant	11	Intermediate
PH31	*Escherichia coli*	6	Resistant	19	Susceptible
PH39	*Escherichia coli*	23	Susceptible	10	Resistant
PH51	*Escherichia coli*	27	Susceptible	6	Resistant
PH5	*Escherichia coli*	21	Susceptible	12	Intermediate
PH85	*Escherichia coli*	24	Susceptible	6	Resistant
PH90	*Escherichia coli*	22	Susceptible	6	Resistant
PH92-1	*Escherichia coli*	24	Susceptible	6	Resistant
PH93	*Escherichia coli*	24	Susceptible	6	Resistant
PH94	*Escherichia coli*	6	Resistant	6	Resistant

PH98	*Escherichia coli*	24	Susceptible	6	Resistant
WU31	*Escherichia coli*	21	Susceptible	9	Resistant
WU32	*Escherichia coli*	17	Susceptible	6	Resistant
WU33	*Escherichia coli*	12	Resistant	20	Susceptible
WU34	*Escherichia coli*	6	Resistant	6	Resistant
WU35	*Escherichia coli*	17	Susceptible	27	Susceptible
WU40	*Escherichia coli*	21	Susceptible	20	Susceptible
WU43	*Escherichia coli*	22	Susceptible	22	Susceptible
WU44	*Escherichia coli*	21	Susceptible	23	Susceptible
WU45	*Escherichia coli*	21	Susceptible	22	Susceptible
PH112-2	*Enterobacter aerogenes*	24	Susceptible	9	Resistant
PH113	*Enterobacter aerogenes*	23	Susceptible	9	Resistant
PH134	*Enterobacter aerogenes*	6	Resistant	12	Intermediate
PH138-2	*Enterobacter aerogenes*	6	Resistant	9	Resistant
PH63	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH84-2	*Enterobacter aerogenes*	6	Resistant	9	Resistant
PH112-1	*Enterobacter cloacae* complex	6	Resistant	16	Susceptible
PH125	*Enterobacter cloacae* complex	6	Resistant	15	Susceptible
PH158	*Enterobacter cloacae* complex	6	Resistant	10	Resistant
PH23	*Enterobacter cloacae* complex	6	Resistant	10	Resistant
PH24-2	*Enterobacter cloacae* complex	6	Resistant	11	Intermediate
PH82	*Enterobacter cloacae* complex	6	Resistant	10	Resistant
WU26	*Enterobacter cloacae* complex	6	Resistant	11	Intermediate
WU27	*Enterobacter cloacae* complex	13	Intermediate	6	Resistant
WU29	*Enterobacter cloacae* complex	21	Susceptible	15	Susceptible
PH102	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH10	*Klebsiella pneumoniae*	6	Resistant	14	Susceptible
PH11	*Klebsiella pneumoniae*	11	Resistant	14	Susceptible
PH124	*Klebsiella pneumoniae*	25	Susceptible	10	Resistant
PH12	*Klebsiella pneumoniae*	6	Resistant	14	Susceptible
PH139	*Klebsiella pneumoniae*	6	Resistant	13	Intermediate

PH150-2	*Klebsiella pneumoniae*	24	Susceptible	6	Resistant
PH152	*Klebsiella pneumoniae*	22	Susceptible	6	Resistant
PH24-1	*Klebsiella pneumoniae*	6	Resistant	14	Susceptible
PH25	*Klebsiella pneumoniae*	6	Resistant	15	Susceptible
PH28-1	*Klebsiella pneumoniae*	22	Susceptible	12	Intermediate
PH38-1	*Klebsiella pneumoniae*	23	Susceptible	12	Intermediate
PH40	*Klebsiella pneumoniae*	24	Susceptible	13	Intermediate
PH44	*Klebsiella pneumoniae*	6	Resistant	15	Susceptible
PH49-2	*Klebsiella pneumoniae*	23	Susceptible	14	Susceptible
PH72	*Klebsiella pneumoniae*	6	Resistant	9	Resistant
PH73	*Klebsiella pneumoniae*	22	Susceptible	17	Susceptible
PH88	*Klebsiella pneumoniae*	6	Resistant	10	Resistant
PH9	*Klebsiella pneumoniae*	6	Resistant	14	Susceptible
WU10	*Klebsiella pneumoniae*	24	Susceptible	18	Susceptible
WU12	*Klebsiella pneumoniae*	22	Susceptible	14	Susceptible
WU18	*Klebsiella pneumoniae*	20	Susceptible	14	Susceptible
WU21	*Klebsiella pneumoniae*	22	Susceptible	6	Resistant
WU23	*Klebsiella pneumoniae*	6	Resistant	10	Resistant
WU2	*Klebsiella pneumoniae*	21	Susceptible	18	Susceptible
WU3	*Klebsiella pneumoniae*	22	Susceptible	19	Susceptible
WU6	*Klebsiella pneumoniae*	24	Susceptible	18	Susceptible
WU7	*Klebsiella pneumoniae*	22	Susceptible	19	Susceptible
WU8	*Klebsiella pneumoniae*	26	Susceptible	11	Intermediate
WU9	*Klebsiella pneumoniae*	23	Susceptible	19	Susceptible
Isolate	Species	DOX ZOI	DOX Interpretation	CHL ZOI	CHL Interpretation
PH100	*Escherichia coli*	6	Resistant	6	Resistant
PH101-2	*Escherichia coli*	6	Resistant	25	Susceptible
PH105	*Escherichia coli*	6	Resistant	6	Resistant
PH108	*Escherichia coli*	6	Resistant	17	Intermediate
PH114	*Escherichia coli*	6	Resistant	6	Resistant
PH118	*Escherichia coli*	6	Resistant	23	Susceptible
PH129	*Escherichia coli*	10	Resistant	23	Susceptible
PH135	*Escherichia coli*	20	Susceptible	27	Susceptible
PH141	*Escherichia coli*	20	Susceptible	27	Susceptible
PH143	*Escherichia coli*	6	Resistant	25	Susceptible
PH151-2	*Escherichia coli*	13	Intermediate	26	Susceptible

PH156-1	*Escherichia coli*	19	Susceptible	6	Resistant
PH18	*Escherichia coli*	11	Intermediate	22	Susceptible
PH20	*Escherichia coli*	11	Intermediate	6	Resistant
PH31	*Escherichia coli*	19	Susceptible	24	Susceptible
PH39	*Escherichia coli*	10	Resistant	25	Susceptible
PH51	*Escherichia coli*	6	Resistant	19	Susceptible
PH5	*Escherichia coli*	12	Intermediate	23	Susceptible
PH85	*Escherichia coli*	6	Resistant	21	Susceptible
PH90	*Escherichia coli*	6	Resistant	25	Susceptible
PH92-1	*Escherichia coli*	6	Resistant	6	Resistant
PH93	*Escherichia coli*	6	Resistant	6	Resistant
PH94	*Escherichia coli*	6	Resistant	25	Susceptible
PH98	*Escherichia coli*	6	Resistant	6	Resistant
WU31	*Escherichia coli*	9	Resistant	6	Resistant
WU32	*Escherichia coli*	6	Resistant	6	Resistant
WU33	*Escherichia coli*	20	Susceptible	6	Resistant
WU34	*Escherichia coli*	6	Resistant	14	Intermediate
WU35	*Escherichia coli*	27	Susceptible	28	Susceptible
WU40	*Escherichia coli*	20	Susceptible	24	Susceptible
WU43	*Escherichia coli*	22	Susceptible	27	Susceptible
WU44	*Escherichia coli*	23	Susceptible	26	Susceptible
WU45	*Escherichia coli*	22	Susceptible	27	Susceptible
PH112-2	*Enterobacter aerogenes*	9	Resistant	6	Resistant
PH113	*Enterobacter aerogenes*	9	Resistant	6	Resistant
PH134	*Enterobacter aerogenes*	12	Intermediate	6	Resistant
PH138-2	*Enterobacter aerogenes*	9	Resistant	6	Resistant
PH63	*Enterobacter aerogenes*	6	Resistant	6	Resistant
PH84-2	*Enterobacter aerogenes*	9	Resistant	6	Resistant
PH112-1	*Enterobacter cloacae* complex	16	Susceptible	6	Resistant
PH125	*Enterobacter cloacae* complex	15	Susceptible	9	Resistant
PH158	*Enterobacter cloacae* complex	10	Resistant	6	Resistant
PH23	*Enterobacter cloacae* complex	10	Resistant	6	Resistant
PH24-2	*Enterobacter cloacae* complex	11	Intermediate	6	Resistant
PH82	*Enterobacter cloacae*	10	Resistant	22	Susceptible

	complex				
WU26	*Enterobacter cloacae* complex	11	Intermediate	17	Intermediate
WU27	*Enterobacter cloacae* complex	6	Resistant	10	Resistant
WU29	*Enterobacter cloacae* complex	15	Susceptible	21	Susceptible
PH102	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH10	*Klebsiella pneumoniae*	14	Susceptible	11	Resistant
PH11	*Klebsiella pneumoniae*	14	Susceptible	12	Resistant
PH124	*Klebsiella pneumoniae*	10	Resistant	16	Intermediate
PH12	*Klebsiella pneumoniae*	14	Susceptible	11	Resistant
PH139	*Klebsiella pneumoniae*	13	Intermediate	6	Resistant
PH150-2	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH152	*Klebsiella pneumoniae*	6	Resistant	6	Resistant
PH24-1	*Klebsiella pneumoniae*	14	Susceptible	12	Resistant
PH25	*Klebsiella pneumoniae*	15	Susceptible	14	Intermediate
PH28-1	*Klebsiella pneumoniae*	12	Intermediate	20	Susceptible
PH38-1	*Klebsiella pneumoniae*	12	Intermediate	19	Susceptible
PH40	*Klebsiella pneumoniae*	13	Intermediate	18	Susceptible
PH44	*Klebsiella pneumoniae*	15	Susceptible	23	Susceptible
PH49-2	*Klebsiella pneumoniae*	14	Susceptible	18	Susceptible
PH72	*Klebsiella pneumoniae*	9	Resistant	6	Resistant
PH73	*Klebsiella pneumoniae*	17	Susceptible	25	Susceptible
PH88	*Klebsiella pneumoniae*	10	Resistant	6	Resistant
PH9	*Klebsiella pneumoniae*	14	Susceptible	11	Resistant
WU10	*Klebsiella pneumoniae*	18	Susceptible	24	Susceptible
WU12	*Klebsiella pneumoniae*	14	Susceptible	6	Resistant
WU18	*Klebsiella pneumoniae*	14	Susceptible	6	Resistant
WU21	*Klebsiella pneumoniae*	6	Resistant	19	Susceptible
WU23	*Klebsiella pneumoniae*	10	Resistant	10	Resistant
WU2	*Klebsiella pneumoniae*	18	Susceptible	24	Susceptible
WU3	*Klebsiella pneumoniae*	19	Susceptible	23	Susceptible
WU6	*Klebsiella pneumoniae*	18	Susceptible	23	Susceptible
WU7	*Klebsiella pneumoniae*	19	Susceptible	24	Susceptible
WU8	*Klebsiella pneumoniae*	11	Intermediate	6	Resistant
WU9	*Klebsiella pneumoniae*	19	Susceptible	26	Susceptible

Table 4.1

In vitro susceptibility testing results.

	Initial ZOI	Initial Interpretation	ZOI A	Interpretation A
WU8	11	Intermediate	10	Resistant
WU26	11	Intermediate	15	Susceptible
PH28-1	12	Intermediate	16	Susceptible
PH38-1	12	Intermediate	16	Susceptible
PH40	13	Intermediate	16	Susceptible
PH139	13	Intermediate	14	Susceptible

	ZOI B	Interpretation B	ZOI C	Interpretation C
WU8	11	Intermediate	11	Intermediate
WU26	15	Susceptible	15	Susceptible
PH28-1	17	Susceptible	17	Susceptible
PH38-1	17	Susceptible	17	Susceptible
PH40	17	Susceptible	17	Susceptible
PH139	14	Susceptible	13	Intermediate

Table 4.2

Retest results of 6 isolates for Doxycycline susceptibility. A, B, and C represent the results of three independent testers in the retest.

Chapter 5: Conclusions

My results from Chapter 2 suggest that HGT happens readily within the *Enterobacteriaceae*, and is a major player in the expansion of KPC and NDM-1 thus far. This fits with our current knowledge of the spread of antibiotic resistance, as other widely disseminated genes, such as CTX-M-15 (31), were originally acquired by pathogens from environmental bacteria. On the other hand, there is a strong phylogenetic bias to successful HGT events (19, 84), and though the human intestinal microbiota is enriched for long distance transfers (19, 20), in Chapter 3 I show that resistance genes are only slightly more likely to have gone through a HGT event detectable by codon bias than genes that are closely tied to phylogeny. This is good news for rapid diagnostics, such as those I described in Chapter 4, since any genotype-based diagnostic will rely on HGT of entirely new resistance genes into pathogens to be rare. This is also good for alternative treatment strategies like combinations or cycling that use existing antibiotics. For example, the combination described in the appendix suppresses the evolution of resistance, but it would likely be ineffective if MRSA were to gain a carbapenemase such as NDM-1 or KPC. When rare long distance antibiotic resistance transfers do happen, though, they will need to be quickly identified, before they become widely disseminated between pathogenic strains and species and are that much more difficult to contain.

Genomics-based AST, if implemented, can serve the additional purpose of rapidly identifying novel resistance factors as they enter the pathogenic population. Then molecular epidemiology for resistance genes could become integrated into the normal clinical pipeline, rather than being an additional task for unusual isolates. This could lead to a more gene-centric approach to epidemiology than the strain specific model currently in place. The machine learning results of Chapter 4 indicate that species identity is not predictive when specific antibiotic

resistance genes are considered separately, lending support to the gene-centric model. Other authors have proposed high resolution phylogenetics to identify specific sub-strains as a potential clinical diagnostic, since certain sub-strains are associated with specific resistance and virulence genes (122). This approach has the advantage of being faster than current methods of WGS, but it would require specific knowledge and a specific test for each important substrain, and the tests may cease to be predictive as new genes enter the population through HGT. On the other hand, a single WGS-based assay can be applied to a greater diversity of pathogens, and it can easily accommodate changes in resistance profiles due to HGT.

Regardless of whether WGS ultimately is used for AST, it will be important to increase our understanding of how pathogen genotype leads to specific resistance phenotypes. An ideal prediction program would be able to look at the entire genome and base its estimate of resistance not only on the presence and identity of resistance genes, but also on the pathogen's regulatory network and metabolic profile, to estimate the expression levels of the resistance genes, antibiotic targets, antibiotic importers, and necessary cofactors. Although these factors are represented by the less than 10% error of predictions based on genes alone, understanding them will require a coordinated effort of genomics, transcriptomics, and proteomics all matched with resistance phenotyping. From a diagnostic perspective alone this line of research would not be worth the time and effort, but achieving that level of understanding over antibiotic resistance may well lead to the development of new combination therapies. By identifying exactly how a bacterium coordinates itself to resist an antibiotic, it should be possible to identify a second compound that would throw that coordination off. This has previously been shown in using specific metabolites to potentiate persister cells to aminoglycoside attack (123), showing that the elements of the combination do not even both need to be antibiotics.

Another field that would benefit from additional basic research is the diversity of conjugative plasmids in the *Enterobacteriaceae*. Though some plasmids spread within bacterial populations with little change to their sequence, in Chapter 2 my comparison of newly and previously sequenced plasmids revealed a wide diversity, but with many shared components. This suggests that plasmid recombination is allowing many different combinations of plasmid components to be tested in pathogens. The most successful plasmid compositions are then likely to expand within the bacterial population and unfortunately, a diversity of antibiotic resistance is likely to determine success, at least in the hospital environment. A true, expansive measure of the diversity of plasmids available to pathogens will allow researchers to make statistical comparisons between different plasmid components, such as those carrying NDM-1 and KPC, potentially allowing predictions of which component will be more successful going into the future. In the case of NDM-1 and KPC this could make a difference in treatment. Though both enzymes can degrade most beta-lactams, KPC is somewhat affected by beta-lactamase inhibitors and is not effective against cephamycins, while NDM-1 is not effective against aztreonam. Under the right circumstances, then, clinicians could exploit those weaknesses and deliver an effective treatment.

Determining the resistance genes that pathogens may get through HGT is clinically important, but it may be equally important to identify resistance genes that pathogens will not acquire. The prevalence of antibiotic resistance in most studied environments can give the impression that attempts to contain highly resistant strains are hopeless, since susceptible pathogens will just receive resistance genes from the environment if not from each other. If many of the resistance genes in the environment are not easily available to pathogens, though, then they will acquire novel resistance only through rare events. That the latter is the case is suggested

not only by my work in Chapter 3, but also by previous work in the soil (114). If such transfer events are indeed rare on human-relevant timescales, then it should be possible to contain dangerous resistances in a specific locations, as is being attempted today with NDM-1 (124). Such containment strategies, especially if used with genotype-based diagnostics and gene-centric epidemiology, could slow the spread of antibiotic resistance in the future, allowing development of new therapeutics to keep pace.

References/Bibliography/Works Cited

1. **CDC.** 2013. Antibiotic Resistance Threats in the United States. Centers for Disease Control and Prevention.,
2. **Fleming SA.** 1945. Penicillin, *on* The Nobel Foundation. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1945/fleming-lecture.html. Accessed 12 Jan 2015.
3. **Lewis K.** 2013. Platforms for antibiotic discovery. Nat Rev Drug Discov **12:**371-387.
4. **D/'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD.** 2011. Antibiotic resistance is ancient. Nature **477:**457-461.
5. **Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, Magris M, Hidalgo G, Contreras M, Noya-Alarcón Ó, Lander O, McDonald J, Cox M, Walter J, Oh PL, Ruiz JF, Rodriguez S, Shen N, Song SJ, Metcalf J, Knight R, Dantas G, Dominguez-Bello MG.** 2015. The microbiome of uncontacted Amerindians. Science Advances **1**.
6. **Perry JA, Westman EL, Wright GD.** 2014. The antibiotic resistome: what's new? Current Opinion in Microbiology **21:**45-50.
7. **Wand ME, Baker KS, Benthall G, McGregor H, McCowen JWI, Deheer-Graham A, Sutton JM.** 2015. Characterisation of pre-antibiotic era Klebsiella pneumoniae isolates with respect to antibiotic/disinfectant susceptibility and virulence in Galleria mellonella. Antimicrobial Agents and Chemotherapy doi:10.1128/aac.05009-14.
8. **Burton DC, Edwards JR, Horan TC, Jernigan JA, Fridkin SK.** 2009. MEthicillin-resistant staphylococcus aureus central line–associated bloodstream infections in us intensive care units, 1997-2007. JAMA **301:**727-736.
9. **Silbergeld EK, Graham J, Price LB.** 2008. Industrial Food Animal Production, Antimicrobial Resistance, and Human Health. Annual Review of Public Health **29:**151-169.
10. **Creech CB, 2nd, Kernodle DS, Alsentzer A, Wilson C, Edwards KM.** 2005. Increasing rates of nasal carriage of methicillin-resistant Staphylococcus aureus in healthy children. Pediatr Infect Dis J **24:**617-621.
11. **Coorevits L, Boelens J, Claeys G.** 2015. Direct susceptibility testing by disk diffusion on clinical samples: a rapid and accurate tool for antibiotic stewardship. European Journal of Clinical Microbiology & Infectious Diseases doi:10.1007/s10096-015-2349-2**:**1-6.
12. **Institute CaLS.** 2013. Performance Standards for antimicrobial susceptibly testing; twenty-third informational supplement, vol 1.
13. **Seitz P, Blokesch M.** 2013. Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. FEMS Microbiology Reviews **37:**336-363.
14. **Balcazar JL.** 2014. Bacteriophages as Vehicles for Antibiotic Resistance Genes in the Environment. PLoS Pathogens **10:**e1004219.
15. **Waldor MK, Mekalanos JJ.** 1996. LysogenicConversionby a Filamentous Phage Encoding Cholera Toxin. Science **272:**1910-1914.
16. **Bonanno L, Loukiadis E, Mariani-Kurkdjian P, Oswald E, Garnier L, Michel V, Auvray F.** 2015. Diversity of Shiga toxin-producing Escherichia coli (STEC) O26:H11:

characterization of stx subtypes and insertion sites of Stx and EspK bacteriophages. Applied and Environmental Microbiology doi:10.1128/aem.00077-15.
17. **Peng Y, Lu J, Wong JJW, Edwards RA, Frost LS, Mark Glover JN.** 2014. Mechanistic Basis of Plasmid-Specific DNA Binding of the F Plasmid Regulatory Protein, TraM. Journal of Molecular Biology **426:**3783-3795.
18. **Cuzon G, Naas T, Truong H, Villegas MV, Wisell KT, Carmeli Y, Gales AC, Venezia SN, Quinn JP, Nordmann P.** 2010. Worldwide diversity of Klebsiella pneumoniae that produce beta-lactamase blaKPC-2 gene. Emerging infectious diseases **16:**1349-1356.
19. **Smillie CS, Smith MB, Friedman J, Cordero OX, David LA, Alm EJ.** 2011. Ecology drives a global network of gene exchange connecting the human microbiome. Nature **480:**241-244.
20. **Caro-Quintero A, Konstantinidis KT.** 2015. Inter-phylum HGT has shaped the metabolism of many mesophilic and anaerobic bacteria. ISME J **9:**958-967.
21. **Beiko R, Ragan M.** 2008. Detecting Lateral Genetic Transfer, p 457-469. *In* Keith J (ed), Bioinformatics, vol 452. Humana Press.
22. **Karlin S, Campbell AM, Mrázek J.** 1998. COMPARATIVE DNA ANALYSIS ACROSS DIVERSE GENOMES. Annual Review of Genetics **32:**185-225.
23. **Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Doré J, Meta HITC, Weissenbach J, Ehrlich SD, Bork P.** 2011. Enterotypes of the human gut microbiome. Nature **473:**174-180.
24. **van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, Visser CE, Kuijper EJ, Bartelsman JFWM, Tijssen JGP, Speelman P, Dijkgraaf MGW, Keller JJ.** 2013. Duodenal Infusion of Donor Feces for Recurrent Clostridium difficile. New England Journal of Medicine **368:**407-415.
25. **Prevention CfDCa.** December 1, 2014 2014. Reports of Selected E. coli Outbreak Investigations, *on* Centers for Disease Control and Prevention. http://www.cdc.gov/ecoli/outbreaks.html. Accessed
26. **Barber AE, Norton JP, Spivak AM, Mulvey MA.** 2013. Urinary Tract Infections: Current and Emerging Management Strategies. Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America **57:**719-724.
27. **Logue CM, Doetkott C, Mangiamele P, Wannemuehler YM, Johnson TJ, Tivendale KA, Li G, Sherwood JS, Nolan LK.** 2012. Genotypic and Phenotypic Traits That Distinguish Neonatal Meningitis-Associated Escherichia coli from Fecal E. coli Isolates of Healthy Human Hosts. Applied and Environmental Microbiology **78:**5824-5830.
28. **Tumbarello M, Trecarichi EM, De Rosa FG, Giannella M, Giacobbe DR, Bassetti M, Losito AR, Bartoletti M, Del Bono V, Corcione S, Maiuro G, Tedeschi S, Celani L, Cardellino CS, Spanu T, Marchese A, Ambretti S, Cauda R, Viscoli C, Viale P.** 2015. Infections caused by KPC-producing Klebsiella pneumoniae: differences in therapy and mortality in a multicentre study. Journal of Antimicrobial Chemotherapy doi:10.1093/jac/dkv086.

29. **Patel PK, Russo TA, Karchmer AW.** 2014. Hypervirulent Klebsiella pneumoniae. Open Forum Infectious Diseases **1:**ofu028.
30. **Yang Y, Bhachech N, Bush K.** 1995. Biochemical comparison of imipenem, meropenem and biapenem: permeability, binding to penicillin-binding proteins, and stability to hydrolysis by β-lactamases. Journal of Antimicrobial Chemotherapy **35:**75-84.
31. **Hawkey PM, Jones AM.** 2009. The changing epidemiology of resistance. The Journal of antimicrobial chemotherapy **64 Suppl 1:**i3-10.
32. **Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schaberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K.** 2015. A new antibiotic kills pathogens without detectable resistance. Nature **517:**455-459.
33. **Fischbach MA, Walsh CT.** 2009. Antibiotics For Emerging Pathogens. Science (New York, NY) **325:**1089-1093.
34. **Walsh C.** 2003. Antibiotics: Actions, Origins, Resistance. ASM Press, Washington, D.C.
35. **Jones RNT, C.; Barry, A.L.; Fuchs, P.C.; Gavan, T.L.; Gerlach E.H.** 1977. Piperacillin (T-1220), a new semisynthetic penicillin: *in vitro* antimicrobial activity comparison with Carbenicillin, Ticarcillin, Ampicillin, Cephalothin, Cefamandole, and Cefoxitin. The Journal of Antibiotics **XXX:**1107-1114.
36. **Barbee LA, Soge OO, Holmes KK, Golden MR.** 2014. In vitro synergy testing of novel antimicrobial combination therapies against Neisseria gonorrhoeae. The Journal of antimicrobial chemotherapy doi:10.1093/jac/dkt540.
37. **Hegreness M, Shoresh N, Damian D, Hartl D, Kishony R.** 2008. Accelerated evolution of resistance in multidrug environments. Proceedings of the National Academy of Sciences of the United States of America **105:**13977-13981.
38. **Lorian V.** 2005. Antibiotics in Laboratory Medicine. Lippincott Williams & Wilkins.
39. **Imamovic L, Sommer MO.** 2013. Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. Science translational medicine **5:**204ra132.
40. **Lazar V, Pal Singh G, Spohn R, Nagy I, Horvath B, Hrtyan M, Busa-Fekete R, Bogos B, Mehi O, Csorgo B, Posfai G, Fekete G, Szappanos B, Kegl B, Papp B, Pal C.** 2013. Bacterial evolution of antibiotic hypersensitivity. Molecular systems biology **9:**700.
41. **Johnson AP, Woodford N.** 2013. Global spread of antibiotic resistance: the example of New Delhi metallo-beta-lactamase (NDM)-mediated carbapenem resistance. Journal of medical microbiology **62:**499-513.
42. **Villa L, Capone A, Fortini D, Dolejska M, Rodriguez I, Taglietti F, De Paolis P, Petrosillo N, Carattoli A.** 2013. Reversion to susceptibility of a carbapenem-resistant clinical isolate of Klebsiella pneumoniae producing KPC-3. The Journal of antimicrobial chemotherapy **68:**2482-2486.
43. **Dominguez EA, Smith TL, Reed E, Sanders CC, Sanders WE, Jr.** 2000. A Pilot Study of Antibiotic Cycling in a Hematology,ÄêOncology Unit¬†,Ä¢,ÄÇ. Infection Control and Hospital Epidemiology **21:**S4-S8.
44. **Gruson D, Hilbert G, Vargas F, Valentino R, Bebear C, Allery A, Bebear C, Gbikpi-Benissan G, Cardinaud J-P.** 2000. Rotation and Restricted Use of Antibiotics in a

Medical Intensive Care Unit. American Journal of Respiratory and Critical Care Medicine **162:**837-843.
45. **Raymond DP, Pelietier, S. J., Crabtree, T.D., Gleason, T.G., Hamm, L.L., Pruett, T.L., Sawyer, R.G.** 2001. Impact of a rotating empiric antibiotic schedule on infectious mortality in an intensive care unit. Critical care medicine **29:**1101-1108.
46. **Shade A, Klimowicz AK, Spear RN, Linske M, Donato JJ, Hogan CS, McManus PS, Handelsman J.** 2013. Streptomycin Application Has No Detectable Effect on Bacterial Community Structure in Apple Orchard Soil. Applied and Environmental Microbiology **79:**6617-6625.
47. **Khadem A, Soler L, Everaert N, Niewold TA.** 2014. Growth promotion in broilers by both oxytetracycline and Macleaya cordata extract is based on their anti-inflammatory properties. British Journal of Nutrition **112:**1110-1118.
48. **Didelot X, Bowden R, Wilson DJ, Peto TEa, Crook DW.** 2012. Transforming clinical microbiology with bacterial genome sequencing. Nature reviews Genetics **13:**601-612.
49. **Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, Suppes R, Feinstein D, Zanotti S, Taiberg L, Gurka D, Cheang M.** 2006. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Critical care medicine **34:**1589-1596.
50. **Gibson MK, Forsberg KJ, Dantas G.** 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J **9:**207-216.
51. **Yigit H, Queenan AM, Anderson GJ, Domenech-Sanchez A, Biddle JW, Steward CD, Alberti S, Bush K, Tenover FC.** 2001. Novel carbapenem-hydrolyzing beta-lactamase, KPC-1, from a carbapenem-resistant strain of Klebsiella pneumoniae. Antimicrobial agents and chemotherapy **45:**1151-1161.
52. **Munoz-Price LS, Poirel L, Bonomo RA, Schwaber MJ, Daikos GL, Cormican M, Cornaglia G, Garau J, Gniadkowski M, Hayden MK, Kumarasamy K, Livermore DM, Maya JJ, Nordmann P, Patel JB, Paterson DL, Pitout J, Villegas MV, Wang H, Woodford N, Quinn JP.** 2013. Clinical epidemiology of the global expansion of Klebsiella pneumoniae carbapenemases. The Lancet Infectious Diseases **13:**785-796.
53. **Castanheira M, Deshpande LM, Mathai D, Bell JM, Jones RN, Mendes RE.** 2011. Early dissemination of NDM-1- and OXA-181-producing Enterobacteriaceae in Indian hospitals: report from the SENTRY Antimicrobial Surveillance Program, 2006-2007. Antimicrob Agents Chemother **55:**1274-1278.
54. **Walsh TR, Weeks, J., Livermore, D. M. & Toleman, M. A.** 2011. Dissemination of NDM-1 positive bacteria in the New Delhi environemnt and its implications for human health: an environemntal point prevalence study. The Lancet Infectious Diseases **11:**355-362.
55. **Dortet L, Poirel L, Nordmann P.** 2014. Worldwide Dissemination of the NDM-Type Carbapenemases in Gram-Negative Bacteria. BioMed research international **2014:**249856.
56. **Thomas CM, Nielsen KM.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nature reviews Microbiology **3:**711-721.
57. **Robledo IE, Aquino EE, Vazquez GJ.** 2011. Detection of the KPC gene in Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, and Acinetobacter baumannii

during a PCR-based nosocomial surveillance study in Puerto Rico. Antimicrob Agents Chemother **55:**2968-2970.

58. **Sartor AL, Raza MW, Abbasi SA, Day KM, Perry JD, Paterson DL, Sidjabat HE.** 2014. Molecular epidemiology of NDM-1 producing Enterobacteriaceae and Acinetobacter baumannii isolates from Pakistan. Antimicrob Agents Chemother doi:10.1128/AAC.02425-14.
59. **Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, Phan MD, Gomes Moriel D, Peters KM, Davies M, Rogers BA, Dougan G, Rodriguez-Bano J, Pascual A, Pitout JD, Upton M, Paterson DL, Walsh TR, Schembri MA, Beatson SA.** 2014. Global dissemination of a multidrug resistant Escherichia coli clone. Proc Natl Acad Sci U S A **111:**5694-5699.
60. **Manji R, Bythrow M, Branda JA, Burnham CAD, Ferraro MJ, Garner OB, Jennemann R, Lewinski MA, Mochon AB, Procop GW, Richter SS, Rychert JA, Sercia L, Westblade LF, Ginocchio CC.** 2014. Multi-center evaluation of the VITEK® MS system for mass spectrometric identification of non-Enterobacteriaceae Gram-negative bacilli. European Journal of Clinical Microbiology & Infectious Diseases **33:**337-346.
61. **Richter SS, Sercia L, Branda JA, Burnham CAD, Bythrow M, Ferraro MJ, Garner OB, Ginocchio CC, Jennemann R, Lewinski MA, Manji R, Mochon AB, Rychert JA, Westblade LF, Procop GW.** 2013. Identification of Enterobacteriaceae by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry using the VITEK MS system. European Journal of Clinical Microbiology & Infectious Diseases **32:**1571-1578.
62. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nature methods **9:**357-359.
63. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome research **18:**821-829.
64. **Borodovsky M, Lomsadze A.** 2011. Gene identification in prokaryotic genomes, phages, metagenomes, and EST sequences with GeneMarkS suite. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] **Chapter 4:**Unit 4.5.1-17.
65. **Eddy SR.** 2009. A new generation of homology search tools based on probabilistic inference. Genome Informatics **23:**205-211.
66. **Angiuoli SV, Salzberg SL.** 2011. Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics (Oxford, England) **27:**334-342.
67. **Talavera G, Castresana J.** 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic biology **56:**564-577.
68. **Stamatakis A.** 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics doi:10.1093/bioinformatics/btu033.
69. **Price MN, Dehal PS, Arkin AP.** 2010. FastTree 2–Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE **5:**e9490.
70. **Bush K, Jacoby GA.** 2010. Updated functional classification of beta-lactamases. Antimicrobial agents and chemotherapy **54:**969-976.
71. **Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T.** 2011. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics **27:**431-432.

72. **Price LB, Johnson JR, Aziz M, Clabots C, Johnston B, Tchesnokova V, Nordstrom L, Billig M, Chattopadhyay S, Stegger M, Andersen PS, Pearson T, Riddell K, Rogers P, Scholes D, Kahl B, Keim P, Sokurenko EV.** 2013. The epidemic of extended-spectrum-beta-lactamase-producing Escherichia coli ST131 is driven by a single highly pathogenic subclone, H30-Rx. mBio **4:**e00377-00313.
73. **Rogers BA, Sidjabat HE, Paterson DL.** 2011. Escherichia coli O25b-ST131: a pandemic, multiresistant, community-associated strain. The Journal of antimicrobial chemotherapy **66:**1-14.
74. **Pena I, Picazo JJ, Rodriguez-Avial C, Rodriguez-Avial I.** 2014. Carbapenemase-producing Enterobacteriaceae in a tertiary hospital in Madrid, Spain: high percentage of colistin resistance among VIM-1-producing Klebsiella pneumoniae ST11 isolates. International journal of antimicrobial agents doi:10.1016/j.ijantimicag.2014.01.021.
75. **Gibson MK FK, Dantas G.** 2014. Improved annotations of antibiotic resistance functions reveals microbial resistomes cluster by ecology. ISME **(In Press)**.
76. **Patel G, Bonomo RA.** 2011. Status report on carbapenemases: challenges and prospects. Expert Review of Anti-infective Therapy **9:**555-570.
77. **Porres-Osante N, Sáenz Y, Somalo S, Torres C.** 2014. Characterization of Beta-lactamases in Faecal Enterobacteriaceae Recovered from Healthy Humans in Spain: Focusing on AmpC Polymorphisms. Microbial Ecology doi:10.1007/s00248-014-0544-9**:**1-9.
78. **Chen SL, Wu M, Henderson JP, Hooton TM, Hibbing ME, Hultgren SJ, Gordon JI.** 2013. Genomic diversity and fitness of E. coli strains recovered from the intestinal and urinary tracts of women with recurrent urinary tract infection. Science translational medicine **5:**184ra160-184ra160.
79. **Bielaszewska M, Rüter C, Kunsmann L, Greune L, Bauwens A, Zhang W, Kuczius T, Kim KS, Mellmann A, Schmidt MA, Karch H.** 2013. Enterohemorrhagic Escherichia coli Hemolysin Employs Outer Membrane Vesicles to Target Mitochondria and Cause Endothelial and Epithelial Apoptosis. PLoS Pathogens **9:**e1003797.
80. **Pesesky MW, Hussain T, Wallace M, Wang B, Andleeb S, Burnham C-AD, Dantas G.** 2015. KPC and NDM-1 Genes in Related Enterobacteriaceae Strains and Plasmids from Pakistan and the United States. Emerging Infectious Disease journal **21**.
81. **Karami N, Martner A, Enne VI, Swerkersson S, Adlerberth I, Wold AE.** 2007. Transfer of an ampicillin resistance gene between two Escherichia coli strains in the bowel microbiota of an infant treated with antibiotics. Journal of Antimicrobial Chemotherapy **60:**1142-1145.
82. **Moran GG, Yehuda C, Mitchell JS, Inna C, Vered S, Shiri N-V.** 2010. Transfer of Carbapenem-Resistant Plasmid from Klebsiella pneumoniae ST258 to Escherichia coli in Patient. Emerging Infectious Disease journal **16:**1014.
83. **van Schaik W.** 2015. The human gut resistome. Philos Trans R Soc Lond B Biol Sci **370**.
84. **Beiko RG, Harlow TJ, Ragan MA.** 2005. Highways of gene sharing in prokaryotes. Proceedings of the National Academy of Sciences of the United States of America **102:**14332-14337.
85. **Waters JL, Salyers AA.** 2013. Regulation of CTnDOT Conjugative Transfer Is a Complex and Highly Coordinated Series of Events. mBio **4**.

86. **Sommer MOa, Dantas G, Church GM.** 2009. Functional characterization of the antibiotic resistance reservoir in the human microflora. Science (New York, NY) **325:**1128-1131.
87. **Nalbantoglu OU, Way SF, Hinrichs SH, Sayood K.** 2011. RAIphy: Phylogenetic classification of metagenomics samples using iterative refinement of relative abundance index profiles. BMC Bioinformatics **12:**41-41.
88. **Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C.** 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. Nature methods **9:**811-814.
89. **Consortium THMP.** 2012. Structure, function and diversity of the healthy human microbiome. Nature **486:**207-214.
90. **Finn RD, Clements J, Eddy SR.** 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Research **39:**W29-W37.
91. **Wu D, Jospin G, Eisen JA.** 2013. Systematic Identification of Gene Families for Use as "Markers" for Phylogenetic and Phylogeny-Driven Ecological Studies of Bacteria and Archaea and Their Major Subgroups. PLoS ONE **8:**e77033.
92. **Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MOa, Dantas G.** 2012. The shared antibiotic resistome of soil bacteria and human pathogens. Science (New York, NY) **337:**1107-1111.
93. **Kothari A, Morgan M, Haake DA.** 2014. Emerging technologies for rapid identification of bloodstream pathogens. Clin Infect Dis **59:**272-278.
94. **Pulido MR, Garcia-Quintanilla M, Martin-Pena R, Cisneros JM, McConnell MJ.** 2013. Progress on the development of rapid methods for antimicrobial susceptibility testing. J Antimicrob Chemother **68:**2710-2717.
95. **Zumla A, Al-Tawfiq JA, Enne VI, Kidd M, Drosten C, Breuer J, Muller MA, Hui D, Maeurer M, Bates M, Mwaba P, Al-Hakeem R, Gray G, Gautret P, Al-Rabeeah AA, Memish ZA, Gant V.** 2014. Rapid point of care diagnostic tests for viral and bacterial respiratory tract infections—needs, advances, and future prospects. The Lancet Infectious Diseases **14:**1123-1135.
96. **Mancini N, Carletti S, Ghidoli N, Cichero P, Burioni R, Clementi M.** 2010. The Era of Molecular and Other Non-Culture-Based Methods in Diagnosis of Sepsis. Clinical Microbiology Reviews **23:**235-251.
97. **Bertelli C, Greub G.** 2013. Rapid bacterial genome sequencing: methods and applications in clinical microbiology. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **19:**803-813.
98. **Wilson MR, Naccache SN, Samayoa E, Biagtan M, Bashir H, Yu G, Salamat SM, Somasekar S, Federman S, Miller S, Sokolic R, Garabedian E, Candotti F, Buckley RH, Reed KD, Meyer TL, Seroogy CM, Galloway R, Henderson SL, Gern JE, DeRisi JL, Chiu CY.** 2014. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. N Engl J Med **370:**2408-2417.
99. **Reuter S, Ellington MJ, Cartwright EJ, Koser CU, Torok ME, Gouliouris T, Harris SR, Brown NM, Holden MT, Quail M, Parkhill J, Smith GP, Bentley SD, Peacock SJ.** 2013. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. JAMA Intern Med **173:**1397-1404.

100. **Snitkin ES, Zelazny AM, Thomas PJ, Stock F, Program NCS, Henderson DK, Palmore TN, Segre JA.** 2012. Tracking a Hospital Outbreak of Carbapenem-Resistant Klebsiella pneumoniae with Whole-Genome Sequencing. Science translational medicine **4:**148ra116-148ra116.
101. **Gordon NC, Price JR, Cole K, Everitt R, Morgan M, Finney J, Kearns AM, Pichon B, Young B, Wilson DJ, Llewelyn MJ, Paul J, Peto TEA, Crook DW, Walker AS, Golubchik T.** 2014. Prediction of Staphylococcus aureus Antimicrobial Resistance by Whole-Genome Sequencing. Journal of Clinical Microbiology **52:**1182-1191.
102. **Hasman H, Saputra D, Sicheritz-Ponten T, Lund O, Svendsen CA, Frimodt-Møller N, Aarestrup FM.** 2014. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. Journal of clinical microbiology **52:**139-146.
103. **Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A.** 2014. Bacterial Whole-Genome Sequencing Revisited: Portable, Scalable, and Standardized Analysis for Typing and Detection of Virulence and Antibiotic Resistance Genes. Journal of Clinical Microbiology **52:**2365-2370.
104. **Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, Johnson JR, Walker aS, Peto TEa, Crook DW.** 2013. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. The Journal of antimicrobial chemotherapy **68:**2234-2244.
105. **Feng Y, Zhang Y, Ying C, Wang D, Du C.** 2015. Nanopore-based Fourth-generation DNA Sequencing Technology. Genomics, Proteomics & Bioinformatics **13:**4-16.
106. **Wetterstrand K.** DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). http://www.genome.gov/sequencingcosts. Accessed April 7, 2015.
107. **Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M.** 2014. Pfam: the protein families database. Nucleic acids research **42:**D222-D230.
108. **Haft DH, Loftus BJ, Richardson DL, Yang F, Eisen Ja, Paulsen IT, White O.** 2001. TIGRFAMs: a protein family resource for the functional identification of proteins. Nucleic acids research **29:**41-43.
109. **Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH.** 2009. The WEKA Data Mining Software: An Update. SIGKDD Explorations **11**.
110. **Molinaro AM, Simon R, Pfeiffer RM.** 2005. Prediction error estimation: a comparison of resampling methods. Bioinformatics **21:**3301-3307.
111. **McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJV, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD.** 2013. The Comprehensive Antibiotic Resistance Database. Antimicrobial Agents and Chemotherapy **57:**3348-3357.
112. **Fu L, Niu B, Zhu Z, Wu S, Li W.** 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics **28:**3150-3152.
113. **Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM.** 2000. Cloning the Soil Metagenome: a

Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. Applied and Environmental Microbiology **66:**2541-2547.
114. **Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N, Dantas G.** 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature **509:**612-616.
115. **George AM, Hall RM, Stokes HW.** 1995. Multidrug resistance in Klebsiella pneumoniae: a novel gene, ramA, confers a multidrug resistance phenotype in Escherichia coli. Microbiology **141:**1909-1920.
116. **Ajbani K, Lin SY, Rodrigues C, Nguyen D, Arroyo F, Kaping J, Jackson L, Garfein RS, Catanzaro D, Eisenach K, Victor TC, Crudu V, Gler MT, Ismail N, Desmond E, Catanzaro A, Rodwell TC.** 2015. Evaluation of pyrosequencing for detecting extensively drug-resistant Mycobacterium tuberculosis among clinical isolates from four high-burden countries. Antimicrob Agents Chemother **59:**414-420.
117. **Koser CU, Bryant JM, Becq J, Torok ME, Ellington MJ, Marti-Renom MA, Carmichael AJ, Parkhill J, Smith GP, Peacock SJ.** 2013. Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. N Engl J Med **369:**290-292.
118. **Bonnet R.** 2004. Growing Group of Extended-Spectrum β-Lactamases: the CTX-M Enzymes. Antimicrobial Agents and Chemotherapy **48:**1-14.
119. **Ruiz J.** 2003. Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection. Journal of Antimicrobial Chemotherapy **51:**1109-1117.
120. **Castanheira M, Mills JC, Farrell DJ, Jones RN.** 2014. Mutation-Driven β-Lactam Resistance Mechanisms among Contemporary Ceftazidime-Nonsusceptible Pseudomonas aeruginosa Isolates from U.S. Hospitals. Antimicrobial Agents and Chemotherapy **58:**6844-6850.
121. **Turton JF, Ward ME, Woodford N, Kaufmann ME, Pike R, Livermore DM, Pitt TL.** 2006. The role of ISAba1 in expression of OXA carbapenemase genes in Acinetobacter baumannii, vol 258.
122. **Johnson JR, Clermont O, Johnston B, Clabots C, Tchesnokova V, Sokurenko E, Junka AF, Maczynska B, Denamur E.** 2014. Rapid and Specific Detection, Molecular Epidemiology, and Experimental Virulence of the O16 Subgroup within Escherichia coli Sequence Type 131. Journal of Clinical Microbiology **52:**1358-1365.
123. **Allison KR, Brynildsen MP, Collins JJ.** 2011. Metabolite-Enabled Eradication of Bacterial Persisters by Aminoglycosides. Nature **473:**216-220.
124. **Guh AY, Limbago BM, Kallen AJ.** 2014. Epidemiology and prevention of carbapenem-resistant Enterobacteriaceae in the United States. Expert Review of Anti-infective Therapy **12:**565-580.
125. **Walsh TR, Weeks J, Livermore DM, Toleman MA.** 2011. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. The Lancet Infectious Diseases doi:10.1016/s1473-3099(11)70059-7.
126. **Davies J.** 2007. Microbes have the last word. EMBO Reports **9:**302-317.
127. **Davies J, Davies D.** 2010. Origins and Evolution of Antibiotic Resistance. Microbiology and Molecular Biology Reviews **74:**417-433.
128. **Fuda CCS, Fisher JF, Mobashery S.** 2005. Beta-lactam resistance in Staphylococcus aureus: the adaptive resistance of a plastic genome. Cellular and molecular life sciences : CMLS **62:**2617-2633.

129. **Chambers HF, Deleo FR.** 2009. Waves of resistance: Staphylococcus aureus in the antibiotic era. Nature reviews Microbiology **7:**629-641.
130. **Malouin F, Bryan LE.** 1986. MINIREVIEWS Modification of Penicillin-Binding Proteins of Beta-Lactam Resistance. Antimicrobial agents and chemotherapy **30:**1-5.
131. **Harrison EM, Paterson GK, Holden MTG, Ba X, Rolo J, Morgan FJE, Pichon B, Kearns A, Zadoks RN, Peacock SJ, Parkhill J, Holmes MA.** 2013. A novel hybrid SCCmec-mecC region in Staphylococcus sciuri. The Journal of antimicrobial chemotherapy doi:10.1093/jac/dkt452**:**1-8.
132. **Fuda C, Suvorov M, Vakulenko SB, Mobashery S.** 2004. The basis for resistance to beta-lactam antibiotics by penicillin-binding protein 2a of methicillin-resistant Staphylococcus aureus. The Journal of biological chemistry **279:**40802-40806.
133. **Fuda C, Hesek D, Lee M, Morio K-i, Nowak T, Mobashery S.** 2005. Activation for Catalysis of Penicillin-Binding Protein 2a from Methicillin-Resistant Staphylococcus aureus by Bacterial Cell Wall. J Am Chem Soc **127:**2056-2057.
134. **Otero LH, Rojas-Altuve A, Llarrull LI, Carrasco-López C, Kumarasiri M.** 2013. How allosteric control of Staphylococcus aureus penicillin binding protein 2a enables methicillin resistance and physiological function. Proceedings of the National Academy of Sciences of the United States of America **110:**16808-16813.
135. **Villegas-Estrada A, Lee M, Hesek D, Vakulenko SB, Mobashery S.** 2008. Co-opting the Cell Wall in Fighting Methicillin-Resistant Staphylococcus aureus : Potent Inhibition of PBP 2a by Two Anti-MRSA β-Lactam Antibiotics. J Am Chem Soc **130:**9212-9213.
136. **Long SW, Olsen RJ, Mehta SC, Palzkill T, Cernoch PL, Perez KK, Musick WL, Rosato AE, Musser JM.** 2014. PBP2a mutations causing high-level Ceftaroline resistance in clinical methicillin-resistant Staphylococcus aureus isolates. Antimicrobial agents and chemotherapy **58:**6668-6674.
137. **Gu B, Kelesidis T, Tsiodras S, Hindler J, Humphries RM.** 2012. The emerging problem of linezolid-resistant Staphylococcus. The Journal of antimicrobial chemotherapy doi:10.1093/jac/dks354.
138. **van Hal SJ, Paterson DL, Gosbell IB.** 2010. Emergence of daptomycin resistance following vancomycin-unresponsive Staphylococcus aureus bacteraemia in a daptomycin-naïve patient-a review of the literature. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology doi:10.1007/s10096-010-1128-3.
139. **Bhusal Y, Shiohira CM, Yamane N.** 2005. Determination of in vitro synergy when three antimicrobial agents are combined against Mycobacterium tuberculosis. International journal of antimicrobial agents **26:**292-297.
140. **Sjolund M, Wreiber K, Andersson DI, Blaser M, Engstrand L.** 2013. Long-Term Persistence of Resistant Enterococcus Species after Antibiotics To Eradicate Helicobacter pylori. Ann Intern Med **139:**483-488.
141. **Tupin A, Gualtieri M, Roquet-Banères F, Morichaud Z, Brodolin K, Leonetti J-P.** 2010. Resistance to rifampicin: at the crossroads between ecological, genomic and medical concerns. International journal of antimicrobial agents **35:**519-523.
142. **Zhong M, Zhang X, Wang Y, Zhang C, Chen G, Hu P, Li M, Zhu B, Zhang W, Zhang Y.** 2010. An interesting case of rifampicin-dependent/-enhanced multidrug-resistant tuberculosis. The international journal of tuberculosis and lung disease : the

official journal of the International Union against Tuberculosis and Lung Disease **14:**40-44.
143. **Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan J, Niemann S, Gagneux S.** 2011. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. Nature Genetics **44:**106-110.
144. **Lázár V, Pal Singh G, Spohn R, Nagy I, Horváth B, Hrtyan M, Busa-Fekete R, Bogos B, Méhi O, Csörgő B, Pósfai G, Fekete G, Szappanos B, Kégl B, Papp B, Pál C.** 2013. Bacterial evolution of antibiotic hypersensitivity. Molecular Systems Biology **9**.
145. **Imamovic L, Sommer MOA.** 2013. Use of Collateral Sensitivity Networks to Design Drug Cycling Protocols That Avoid Resistance Development. Science Translational Medicine **5:**204ra132-204ra132.
146. **Zimmermann GR, Lehár J, Keith CT.** 2007. Multi-target therapeutics: when the whole is greater than the sum of the parts. Drug discovery today **12:**34-42.
147. **Hegreness M, Shoresh N, Damian D, Hartl DL, Kishony R.** 2008. Accelerated evolution of resistance in multidrug environments. Proceedings of the National Academy of Sciences of the United States of America **105:**13977-13981.
148. **Boucher HW, Talbot GH, Bradley JS, Edwards JE, Gilbert D, Rice LB, Scheld M, Spellberg B, Bartlett J.** 2009. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **48:**1-12.
149. **Rice LB.** 2006. Antimicrobial resistance in gram-positive bacteria. American journal of infection control **34:**S11-19; discussion S64-73.
150. **Alekshun MN, Levy SB.** 2007. Molecular mechanisms of antibacterial multidrug resistance. Cell **128:**1037-1050.
151. **Kumarasamy KK, Toleman MA, Walsh TR, Bagaria J, Butt F, Balakrishnan R, Chaudhary U, Doumith M, Giske CG, Irfan S.** 2010. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. The Lancet Infectious Diseases **3099:**1-6.
152. **Spratt BG.** 1977. Properties of the penicillin-binding proteins of Escherichia coli K12. European journal of biochemistry / FEBS **72:**341-352.
153. **Waxman DJ, Strominger JL.** 1983. Penicillin-binding proteins and the mechanism of action of beta-lactam antibiotics. Annual review of biochemistry **52:**825-869.
154. **Lee SH, Jarantow LW, Wang H, Sillaots S, Cheng H, Meredith TC, Thompson J, Roemer T.** 2011. Antagonism of chemical genetic interaction networks resensitize MRSA to β-lactam antibiotics. Chemistry & biology **18:**1379-1389.
155. **Koga T, Sugihara C, Kakuta M, Masuda N, Namba E, Fukuoka T.** 2009. Affinity of Tomopenem (CS-023) for penicillin-binding proteins in Staphylococcus aureus, Escherichia coli, and Pseudomonas aeruginosa. Antimicrobial agents and chemotherapy **53:**1238-1241.
156. **Yang Y, Bhachech N, Bush K.** 1995. Biochemical comparison of imipenem, meropenem and biapenem: permeability, binding to penicillin-binding proteins, and stability to hydrolysis by beta-lactamases. The Journal of antimicrobial chemotherapy **35:**75-84.

157. **Campbell EM, Chao L.** 2014. A population model evaluating the consequences of the evolution of double-resistance and tradeoffs on the benefits of two-drug antibiotic treatments. PloS one **9:**e86971-e86971.
158. **Munck C, Gumpert HK, Wallin AIN, Wang HH, Sommer MOA.** 2014. Prediction of resistance development against drug combinations by collateral responses to component drugs. Science Translational Medicine **6:**262ra156-262ra156.
159. **Clinical, Laboratory Standards I.** 2009. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically ; Approved Standard — Eighth Edition. CLSI **29**.
160. **Cai Y, Wang R, Pei F, Liang B-b.** 2007. Antibacterial Activity of Allicin Alone and in Combination with Beta-Lactams against Staphylococcus spp. and Pseudomonas aeruginosa. Journal of Antibiotics **60:**335-338.
161. **Berenbaum MC.** 1978. A method for testing for synergy with any number of agents. The Journal of infectious diseases **137:**122-130.
162. **Clinical, Laboratory Standards I.** 1999. Methods for Determining Bactericidal Activity of Antimicrobial Agents; Approved Guideline. CLSI **19**.
163. **Life_Technologies.** 2012. SYBR® Select Master Mix for CFX User Guide.
164. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078-2079.
165. **Somani P, Freimer EH, Gross ML, Higgins JT.** 1988. Pharmacokinetics of Imipenem-Cilastatin in Patients with Renal Insufficiency Undergoing Continuous Ambulatory Peritoneal Dialysis. Antimicrobial agents and chemotherapy **32:**4-9.
166. **Kinzig M, Brismar B, Nord CE.** 1997. Pharmacokinetics and Tissue Penetration of Tazobactam and Piperacillin in Patients Undergoing Colorectal Surgery. Antimicrobial agents and chemotherapy **36:**1997-2004.
167. **Pfizer.** 2012. Zosyn (Piperacillin and Tazobactam for Injection, USP) for injection for intravenous use only prescribing information. Philadelphia, PA.
168. **AstraZeneca Pharmaceuticals LP.** 2013. Merrem IV (meropenem) for injection for intravenous use only prescribing information. Wilmington, DE.
169. **Merck.** 2014. PRIMAXIN I.V. (IMIPENEM AND CILASTATIN FOR INJECTION). Whitehouse Station, NJ.
170. **Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama a, Murakami H, Hosoyama a, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita a, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K.** 2001. Whole genome sequencing of meticillin-resistant Staphylococcus aureus. Lancet **357:**1225-1240.
171. **Goldstein F, Perutka J, Cuirolo A, Plata K, Faccone D, Morris J, Sournia A, Kitzis MD, Ly A, Archer G, Rosato AE.** 2007. Identification and phenotypic characterization of a β-lactam-dependent, methicillin-resistant Staphylococcus aureus strain. Antimicrobial agents and chemotherapy **51:**2514-2522.
172. **Arêde P, Ministro J, Oliveira DC.** 2013. Redefining the role of the β-lactamase locus in methicillin-resistant Staphylococcus aureus: β-lactamase regulators disrupt the MecI-mediated strong repression on mecA and optimize the phenotypic expression of

resistance in strains with constitutive mecA. Antimicrobial agents and chemotherapy **57:**3037-3045.
173. **Berenbaum MC.** 1989. What is Synergy? Pharmacological Reviews **1989**.
174. **Saiman L.** 2007. Clinical utility of synergy testing for multidrug-resistant Pseudomonas aeruginosa isolated from patients with cystic fibrosis: 'the motion for'. Paediatric respiratory reviews **8:**249-255.
175. **Clinical, Laboratory Standards I.** 2009. Performance Standards for Antimicrobial Susceptibility Testing ; Nineteenth Informational Supplement. CLSI **29**.
176. **Fishovitz J, Rojas-Altuve A, Otero LH, Dawley M, Carrasco-lo C, Chang M, Hermoso JA, Mobashery S.** 2014. Disruption of Allosteric Response as an Unprecedented Mechanism of Resistance to Antibiotics. Journal of the American Chemical Society **136:**9814-9817.
177. **Stutman HR, Welch DF, Scribner RK, Marks MI.** 1984. In vitro antimicrobial activity of aztreonam alone and in combination against bacterial isolates from pediatric patients. Antimicrobial agents and chemotherapy **25:**212-215.
178. **Lee SH, Jarantow LW, Wang H, Sillaots S, Cheng H, Meredith TC, Thompson J, Roemer T.** 2011. Antagonism of chemical genetic interaction networks resensitize MRSA to beta-lactam antibiotics. Chem Biol **18:**1379-1389.
179. **Ankomah P, Johnson PJT, Levin BR.** 2013. The pharmaco -, population and evolutionary dynamics of multi-drug therapy: experiments with S. aureus and E. coli and computer simulations. PLoS Pathogens **9:**e1003300-e1003300.
180. **Sonstein SA, Baldwin JN.** 1972. Loss of the penicillinase plasmid after treatment of Staphylococcus aureus with sodium dodecyl sulfate. Journal of bacteriology **109:**262-265.
181. **Hackbarth CJ, Chambers HF.** 1993. blaI and blaR1 regulate beta-lactamase and PBP 2a production in methicillin-resistant Staphylococcus aureus. Antimicrobial agents and chemotherapy **37:**1144-1149.
182. **Lowy FD.** 2003. Antimicrobial resistance : the example of Staphylococcus aureus. Journal of Clinical Investigation **111:**1265-1273.
183. **Blazquez B, Llarrull LI, Luque-ortega JR, Alfonso C, Boggess B, Mobashery S.** 2014. Regulation of the Expression of the β - Lactam Antibiotic-Resistance Determinants in Methicillin-Resistant Staphylococcus aureus (MRSA). Biochemistry **53:**1548-1550.
184. **Arias CA, Murray BE.** 2009. Antibiotic-resistant bugs in the 21st century--a clinical super-challenge. The New England journal of medicine **360:**439-443.
185. **Craig WA.** 1997. The pharmacology of meropenem, a new carbapenem antibiotic. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America **24 Suppl 2:**S266-275.
186. **DeRyke CA, Banevicius MA, Fan HW, Nicolau DP.** 2007. Bactericidal activities of meropenem and ertapenem against extended-spectrum-β-lactamase-producing Escherichia coli and Klebsiella pneumoniae in a neutropenic mouse thigh model. Antimicrobial Agents and Chemotherapy **51:**1481-1486.

Appendix 1: Synergistic, collaterally sensitive triple β-lactam combinations suppress evolution of resistance in Methicillin-Resistant *Staphylococcus aureus* (MRSA)

Patrick R. Gonzales[1], Mitchell W. Pesesky[1], Renee Bouley[2], Anna Ballard[1], Brent A. Biddy[1], Mark A. Suckow[3], William R. Wolter[3], Valerie A. Schroeder[3], Carey-Ann D. Burnham[4,5], Shahriar Mobashery[2], Mayland Chang[2], Gautam Dantas[1,4,6]

[1]Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, Missouri 63108, USA

[2]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, USA

[3]Freimann Life Sciences Center and Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556, USA

[4]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110, USA

[5]Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri 63110, USA

[6]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, Missouri 63130, USA

This work was a collaborative effort involving all of the authors. P.R.G. performed all synergy, collateral sensitivity, and adaptive evolution and was the primary author of the text. M.W.P. performed all sequence and statistical analysis and generated the figures. R.B. performed knockdown experiments and contributed to the *in vivo* experiment, along with M.A.S., W.R.W., and V.A.S. A.B. Assisted P.R.G. with the collateral sensitivity experiments, and B.A.B. assisted P.R.G. with DNA isolation from the adaptive evolution experiment. C-A.D.B. contributed the MRSA isolate collection and clinical expertise to the writing. S.M., M.C., G.D., and P.R.G. designed the experiments.

A1.1 Abstract

Methicillin-resistant *Staphylococcus aureus* (MRSA) is one of the most prevalent multidrug-resistant pathogens worldwide, exhibiting increasing resistance to the latest antibiotic monotherapies used to treat these infections. Here we show that the triple β-lactam combination meropenem/piperacillin/tazobactam (ME/PI/TZ) acts synergistically and is bactericidal against MRSA N315 and 72 clinical MRSA isolates *in vitro*, and clears MRSA N315 infection in a mouse model. ME/PI/TZ suppresses evolution of resistance in MRSA via reciprocal collateral sensitivity of its constituents. We demonstrate that these activities extend to other carbapenem/penicillin/β-lactamase inhibitor combinations. ME/PI/TZ circumvents the tight regulation of the *mec* and *bla* operons in MRSA, the basis for inducible resistance to β-lactam antibiotics. Furthermore, ME/PI/TZ subverts the function of penicillin-binding protein 2a (PBP2a) action via allostery, which we propose as the mechanism for both synergy and collateral sensitivity in this system. Showing similar *in vivo* activity to linezolid, ME/PI/TZ demonstrates that older β-lactam antibiotics could be effective against MRSA infections in humans.

A1.2 Introduction

Multidrug-resistant (MDR) pathogens represent a growing threat to human health, with many infectious diseases effectively regressing toward the pre-antibiotic era (125-127), exemplified by the dramatic rise of community-acquired methicillin-resistant *Staphylococcus aureus* (MRSA) infections. In the 1940's, *S. aureus* infections were primarily treated with first-generation β-lactams (penicillins) which target the penicillin-binding proteins (PBPs), the critical transpeptidases for cell-wall synthesis (128). Four PBPs (PBP1-PBP4) perform these functions in *S. aureus* (128). Emergence of β-lactamase-producing strains led to development of β-lactamase-resistant second-generation penicillins, including methicillin. Soon after the introduction of

methicillin in 1959, the first MRSA strains were reported (129). These strains acquired a highly regulated collection of genes from a non-*S. aureus* source that produced inducible resistance to β-lactam antibiotics (128). One of these genes, *mecA*, encodes penicillin-binding protein 2a (PBP2a). PBP2a performs the critical transpeptidase reaction that cross-links the cell wall, even under challenge by β-lactam antibiotics, when other PBPs are inhibited (130-132). The mechanistic basis for this outcome is complex, involving a closed conformation for the active site, whose function is regulated by allostery (133, 134). The emergence of MRSA has virtually eliminated the use of β-lactams as therapeutic options against *S. aureus*. The recently developed β-lactam agent ceftaroline, which exhibits activity in treatment of MRSA infections, does so by binding to the allosteric site of PBP2a, triggering opening of the active site for inactivation by the drug (134, 135); however, resistance to ceftaroline (136) and other antibiotics used to treat MRSA, including linezolid, vancomycin, and daptomycin, has been reported (137, 138).

Use of higher-order combination therapy targeting orthogonal cellular processes has been successful in treating *Mycobacterium tuberculosis*, *Helicobacter pylori*, and other infections (139, 140). However, resistance is increasing even against these therapies (141-143). We have identified a new potential therapy against MRSA consisting of a combination of clinically approved drugs from three distinct generations and subclasses of β-lactam antibiotics, all targeting cell-wall synthesis: meropenem, piperacillin, and tazobactam (ME/PI/TZ). This therapy uses elements from three strategies: 1) use of semi-synthetic β-lactam antibiotics (33, 136), 2) collateral sensitivity (144, 145), and 3) combinations that increase drug potency by utilizing drug synergy (146, 147). Each of these methods have been successfully employed against the major MDR Gram-negative and Gram-positive human pathogens (148, 149). However, used

individually these strategies have often been thwarted by the evolution of new resistance in MDR pathogens, leading to diminishing options for treating their infections (33, 129, 138, 150, 151).

We hypothesize that ME/PI/TZ operates through inhibition of PBP1 by meropenem, the targeting of PBP2 by piperacillin, protection of piperacillin from the PC1 class A β-lactamases by tazobactam (130, 152-156), and allosteric opening of the active site of PBP2a by meropenem for inhibition by another molecule of antibiotic in the combination (135). This culminates in a synergistic response by simultaneous perturbation of multiple components of the cell-wall synthesis machinery in MRSA. We find that exposure of MRSA N315 to the components of ME/PI/TZ reveals reciprocal collateral sensitivities within this highly synergistic triple combination that suppress the evolution of resistance, in contrast to some synergistic combination therapies that instead accelerate resistance evolution (147, 157). This effect is consistent with recent work showing that collateral sensitivity slows evolution of resistance in a non-pathogenic laboratory strain of *Escherichia coli* (144, 158). Our results support renewed clinical use of older β-lactam antibiotics against MRSA when used in judiciously conceived synergistic combinations of collaterally sensitive components, opening a new treatment paradigm with existing drugs that are already approved for human use.

A1.3 Methods and Materials

Microbiological Studies

MRSA N315 was a gift from Dr. Steven Gill, University of Rochester, Rochester, NY, USA. *S. aureus* ATCC 29213 was acquired from the American Type Culture Collection. De-identified clinical MRSA isolates were selected at random from the clinical isolate strain bank at Barnes-Jewish Hospital, St. Louis, MO, USA. Minimal inhibitory concentration (MIC) assays

for inhibition of growth were performed following the recommendations of the Clinical and Laboratory Standards Institute (CLSI) (159). Briefly, 23 antibacterial compounds (Table A1.1) were selected based on coverage of all major drug classes, including three compounds not classified as antibiotics for human use, but with known antibacterial properties. Compounds were dissolved in dimethyl sulfoxide (DMSO) to a stock concentration of 50 mg/ml. Exceptions: Sulfometuron at 20 mg/ml in DMSO; Tobramycin, D-cycloserine, and colistin at 50 mg/ml in H_2O and filtered at 2 µm. The 23 compounds were formulated into all 253 possible unique pairwise combinations at fixed ratios and at 100x concentrations in solvent. To increase the range of concentrations assayed for possible synergistic or antagonistic drug interactions (>2,000-fold), the drug stocks were arrayed into threefold dilution series down eight rows in 96-well Costar master drug plates, using a BioMek FX robotic liquid handler (Beckman Coulter, Inc.). Drugs were then mixed 1:100 into 96-well plates containing 200 µl/well of cation-adjusted Mueller-Hinton broth (CAMHB). All drug susceptibility assay wells were inoculated with ~1 µl of mid-log phase bacterial culture at 0.5 McFarland standard (~2×10^8 CFU/ml) and grown at 37 °C for 24 h. Endpoint growth at 37 °C after 24 h was determined by optical density at 600 nm ≥0.1 using a Synergy H1 reader (BioTek, Inc.).

Synergy of antibiotic combinations was determined using the fractional inhibitory concentration index (FICI) method (160, 161). By this method, the MIC of the antibiotic compound in combination is divided by the MIC of the compound alone, yielding the fractional contribution of each drug component in the combination. Quotients for all compounds in a combination are summed and drug interactions scored using the formula:

$$FICI = \left(MIC\ A_{comb\ A/B/C}\right)/MIC_{agent\ A} + \left(MIC\ B_{comb\ A/B/C}\right)/MIC_{agent\ B} + \left(MIC\ C_{comb\ A/B/C}\right)/MIC_{agent\ C}$$

Select pairwise combinations against MRSA were then combined with each of the 21 remaining single drugs to make triple combinations, formulated and tested in identical fashion to the double combinations. Synergy of combinations was confirmed via triplicate measurements of drug conditions at the MIC. Based on its high synergy against MRSA N315 in the sparse screening, ME/PI/TZ and its constituents were selected for further characterization. Final susceptibility testing of ME/PI/TZ and its components was performed using twofold dilution from 128 to 2 μg/ml for each component.

Minimal bactericidal concentration (MBC) for ME/PI/TZ in MRSA N315 was determined via duplicate wells of ME/PI/TZ at indicated concentrations in CAMHB media, inoculated with ~5×10^5 CFU/ml of MRSA N315 in mid-log phase and incubated at 37 °C for 24 h. 100 μl of a 1:100 dilution of 50 μl drawn from duplicate ME/PI/TZ wells was plated on Mueller-Hinton agar (MHA) plates and incubated overnight for 24 h. No colony growth at or two dilutions above the MIC confirmed bactericidal activity, as defined by CLSI (162). Meropenem (CAS 96036-03-2) and clavulanate (CAS 61177-45-5) were obtained from AK Scientific, Inc. (Union City, CA, USA). Piperacillin (CAS 59703-84-3), tazobactam (CAS 89786-04-9), imipenem (CAS 74431-23-5), and amoxicillin (CAS 26787-78-0) were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA).

Adaptation and Cross-resistance Assays

MRSA N315 was grown in 150 μl/well of CAMHB with constant shaking at 37 °C and passaged over 11 days in identical 96-well plates containing replicate threefold dilutions of ME/PI/TZ, ME/PI, ME/TZ, PI/TZ, ME, PI, and TZ. Top concentrations of drug combinations were 33.3 μg/ml for each component, while top concentrations for single drugs was 100 μg/ml. To test for cell viability, at the end of the assay on day 11, all wells from the plate were pinned with a sterile

96-pin replicator and transferred to CAMHB only. After passage, plates were filled 1:1 with 30% CAMHB/glycerol and frozen at -80 °C for later analysis.

Growth rate of isolates over passages in each condition was determined by linear best-fit of logarithm-converted exponential growth phase. Following Hegreness et al (147), those wells containing cells in drug conditions whose growth rates were >0.2 h^{-1} between day one and the average of the last six days of growth were considered significantly adapting to conditions and an adaptation rate α was generated. Adapted isolates were retrospectively chosen from each combination or single compound in wells showing an increase in MIC or growth rate, frozen isolates were streaked out on agar plates to obtain single colonies, re-grown in broth conditions identical to those in which they grew originally, and then re-inoculated in sterile 96-well plates identical to the original 11-day plates.

Expression profiling with qRT-PCR

Wild-type and adapted MRSA N315 isolates were grown in triplicate in 100 ml flasks to mid-log phase in CAMHB +/- piperacillin at 11.1 μg/ml or tazobactam at 33.3 μg/ml. To harvest cells at mid-log phase, each culture flask was split into 2 x 50 ml screw-cap tubes, spun down at 4 ˚C for 10 min at 3500 rpm, supernatant removed, and pellets combined carefully with a 2 ml serological pipette. 1 ml RNAprotect Bacteria Reagent (Qiagen, Valencia, CA, USA) was added to pellets to stabilize the RNA, vortexed briefly, and incubated for 5 min at RT. After incubation, tubes were spun again at 4 ˚C for 10 min at 3500 rpm, supernatant removed, and the pellets were stored at -80 ˚C. Total RNA was extracted by the following protocol:

(1) Resuspend cell pellets in 500 μl Buffer B (200 mM NaCl, 20 mM EDTA).

(2) Add 210 μl 20% SDS.

(3) Add ~250 µl volume of acid-washed sterile glass beads (Sigma, Inc.).

(4) Add 500 µl Phenol:Chloroform:IAA.

(5) Bead beat on 'high' for 5 min.

(6) Spin at 8000 rpm at 4˚C for 3 min (to separate the phases).

(7) Remove top aqueous phase and transfer into a new tube.

(8) Add 700 µl isopropanol.

(9) Add 70 µl 3M NaOAc, mix thoroughly by inversion.

(10) Spin at 4˚C, max rpm, for 10 min.

(11) Aspirate supernatant.

(12) Add 750 µl ice cold 70% EtOH, spin at max rpm at 4˚C for 5 min.

(13) Aspirate supernatant, to let the EtOH dry, leave tubes open in RNase free area.

(14) Add 100 µl nuclease free water to each tube and resuspend (put tubes in 50 ˚C heat block, vortexing periodically).

(15) Add 12 µl TURBO-DNase buffer (Ambion, Inc.) and 10 µl RNase-free TURBO-DNase to each sample, and incubate at 37 ˚C for 30 min.

(16) Purify samples using MEGAClear columns and kit per manufacturer protocol.

(17) Re-purify samples using Baseline-ZERO DNase buffer (Epicentre, Inc.) and 10 µl Baseline-ZERO DNase, following manufacturer protocol.

(18) Elute final RNA samples with 30 µl TE buffer, pH 7.0.

First-strand cDNA was synthesized from total RNA with SuperScript First-Strand Synthesis System for RT-PCR (Life Technologies, Carlsbad, CA, USA). qRT-PCR of *pbp2*, *mecA* and *blaZ* in MRSA N315 was performed against *gyrB* using SYBR Select Master Mix for CFX (Life Technologies, Carlsbad, CA, USA) on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, Inc, Hercules, CA, USA). Primer sequences used (0.3 μM each): pbp2_F: CGTGCCGAAATCAATGAAAGACGC, pbp2_R: GGCACCTTCAGAACCAAATCCACC; mecA_F: TGGAACGATGCCTATCTCATATGC, mecA_R: CAGGAATGCAGAAAGACCAAAGC; blaZ_F: TTTATCAGCAACCTTATAGTCTTTTGGAAC, blaZ_R: CCTGCTGCTTTCGGCAAGAC, gyrB_F: CGATGTGGATGGAGCGCATATTAG, gyrB_R: ACAACGGTGGCTGTGCAATATAC. CFX protocol: 2 min @ 50 ˚C, 2 min @ 95 ˚C, (15 s @ 95 ˚C, 1 min @ 60 ˚C) x 40 cycles. Gene expression was determined using the *ΔΔCt* method of normalized quantitation (163), where *Ct* indicates the cycle number at which exponential growth phase increases above threshold fluorescence signal.

Sequencing library preparation

Genomic DNA (gDNA) was extracted from wild-type and adapted MRSA N315 using lysostaphin digestion and phenol:chloroform:IAA extraction as follows:

(1) Draw 1 ml aliquots from overnight 5 ml shaking cultures of *S. aureus* strains, spin down at 13,000 rpm for 3 min, pour off media, add additional 1 ml of culture and repeat.

(2) Add 500 μl of 2X Buffer A (NaCl 200 mM, Tris 200 mM, EDTA 20 mM) at 4 °C to pelleted cells and vortex briefly to resuspend cells.

(3) Add 2.5 μl of 10 mg/ml (200x) lysostaphin (Sigma-Aldrich, Inc.) to tubes.

(4) Flick mix and spin down tubes, place in 37 ˚C dry bath for 1 h.

(5) Fast cool micro-centrifuge to 4 ˚C.

(6) Add ~250 µl of 0.1 mm zirconium beads (BioSpec Products, cat# 1107910).

(7) Add 210 µl of 20% SDS.

(8) Add 500 µl phenol:chloroform:IAA (25:24:1, pH 7.9), chill samples on ice.

(9) Bead beat on the "homogenize" setting for 4 min (beat 2 min, ice 2 min, beat 2 min).

(10) Spin at 6800 rcf (4 ˚C) for 3min.

(11) Spin down PLG columns (5Prime, cat#2302820) at max speed (20,800 rcf) for 30 s at RT while waiting.

(12) Transfer aqueous phase (~500 µl) to pre-spun phase-lock gel tube.

(13) Add equal amount (500 µl) of phenol:chloroform:IAA (25:24:1, pH 7.9) to tube and mix by inversion (DO NOT VORTEX).

(14) Spin tubes at max speed (20,800 rcf) (RT) for 5 min.

(15) Transfer aqueous phase (~500 µl) to a new Eppendorf tube.

(16) Add 500 µl of -20 ˚C isopropanol.

(17) Add 50 µl (1/10 vol.) of 3M NaOAc at pH 5.5 (Ambion, AM9740), and mix thoroughly by inversion.

(18) Store at -20 ˚C for at least 1h (overnight is preferable but not necessary).

(19) Spin at max speed at 4 ˚C for 20 min.

(20) Wash pellet with 500 µl of 100% EtOH (RT) and spin down at 4 ˚C for 3 min.

(21) Carefully pipet off EtOH, air-dry >15 min.

(22) Add 30 µl of TE (Ambion, AM 9861), incubate at 50 ˚C for 5 min.

(23) Run DNA through QIAGEN QIAQuick PCR purification column with the following modifications: RNase A treatment at beginning of column clean-up. Combine 4 µl Qiagen RNase (100 mg/ml) with every 300 µl buffer PB used, incubate in buffer PB/RNase for 15 min at RT.

(24) Let PE wash buffer sit in column at RT for 2 min, elute gDNA with 35 µl of EB buffer pre-heated to 55 ˚C, letting sit for 1 min before final spin.

We sheared 500 ng of total DNA from each genome to ~300 bp fragments in nine rounds of shearing of ten min each on the BioRuptor XL. In each round the power setting was 'H' and samples were treated for 30 s and allowed to rest for 30 s. Each sample was concentrated using the Qiagen MinElute PCR purification kit per the manufacturer's protocol. End Repair of the sheared DNA fragments was initiated with the addition of 2.5 µl of T4 DNA ligase buffer with 10 mM ATP (NEB, B0202S), 1 µl of 1 mM dNTPs (NEB), 0.5 µl T4 Polymerase (NEB, M0203S), 0.5 µl T4 PNK (NEB M0201S), and 0.5 µl Taq Polymerase (NEB, M0267S). This mixture was incubated at 25 °C for 30 min, then at 75 °C for 20 min. Barcoded adapters were then added to the solution along with 0.8 µl of T4 DNA ligase (NEB, M0202M), for the purpose of ligating the adapters to the DNA fragments. This solution was then incubated at 16 ºC for 40

min, then 65 °C for 10 min. The adapter-ligated DNA was then purified using the Qiagen MinElute PCR purification kit per the manufacturer's protocol.

The DNA fragments were then size selected on a 2% agarose gel in 1X TBE buffer stained with Biotium GelGreen dye (Biotium). DNA fragments were combined with 2.5 µl 6x Orange loading dye before loading on to the gel. Adapter-ligated DNA was extracted from gel slices corresponding to DNA of 250-300 bp using a QIAGEN MinElute Gel Extraction kit per the manufacturer's protocol. The purified DNA was enriched by PCR using 12.5 µl 2x Phusion HF Master Mix and 1 µl of 10 µM Illumina PCR Primer Mix in a 25 µl reaction using 1 µl of purified DNA as template. DNA was amplified at 98 °C for 30 s followed by 18 cycles of 98 °C for 10 s, 65 °C for 30 s, 72°C for 30 s with a final extension of 5 min at 72 °C. The DNA concentration was then measured using the Qubit fluorometer and 10 nmol of each sample (up to 106 samples per lane of sequencing) were pooled. Subsequently, samples were submitted for Illumina HiSeq-2500 Paired-End (PE) 101 bp sequencing at GTAC (Genome Technology Access Center, Washington University in St. Louis) at 9 pmol per lane.

DNA Sequence analysis

Alignment and variant calling. For the wild-type and adapted MRSA N315, all sequencing reads for each genome were de-multiplexed by barcode into separate genome bins. Reads were quality trimmed to remove adapter sequence and bases on either end with a quality score below 19. Any reads shorter than 31 bp after quality trimming were not used in further analysis. All reads were mapped to the *Staphylococcus aureus* subsp. aureus N315 chromosome (GenBank ID: BA000018.3) and pN315 plasmid (GenBank ID: AP003139) (command: bowtie2 –x <reference_genome_index_name> -1 <forward_read_file> -2 <reverse_read_file> -q --phred33 --very-sensitive-local -I 200 -X 1000 -S <sam_output>). Variants from the reference were called

using samtools (164) (commands: samtools view -buS <sam_file> | samtools sort -m 4000000000 - <sample_prefix> ### samtools index <bam_file> ### samtools mpileup -uD -f <reference_genome> <bam_file> | bcftools view -bcv - > <bcf_file> ### bcftools view <bcf_file>). The variant call format (VCF) file was then filtered to remove SNPs with a quality score lower than 70 or coverage greater than twice the average coverage expected per base. Absence of read coverage or overabundant read coverage indicated plasmid loss or large duplication respectively. Any variant position found from the wild-type alignment was determined to be a result of alignment error or to be derived from lab specific drift in N315 and was removed from all other VCF files. Each variant position was then compared to known ORF locations in N315 to search for causal variants.

In vivo mouse model of MRSA infection

Animals. Outbred ICR female mice (6-8 weeks old, 17-25 g body weight; Harlan Laboratories, Inc., Indianapolis, IN, USA) were used. Mice were given Teklad 2019 Extruded Rodent Diet (Harlan Laboratories, Inc., Indianapolis, IN, USA) and water *ad libitum*. Mice were maintained in polycarbonate shoebox cages containing corncob (The Andersons, Inc., Maumee, OH, USA) and Alpha-dri (Shepherd Specialty Papers, Inc., Richland, MI, USA) bedding under 12-h light/12-h dark cycle at 22 ± 1 °C. All procedures involving animals were approved by the University of Notre Dame Institutional Animal Care and Use Committee.

Neutropenic mouse peritonitis model of MRSA infection. Doses of cyclophosphamide (100 µl of 50 mg/ml in 0.9% saline corresponding to 200 mg/kg; Alfa Aesar, Ward Hill, MA, USA) were given intraperitoneally (IP) at 4 days and 1 day prior to infection. The *S. aureus* strain N315 was streaked onto Brain-Heart Infusion (BHI; Becton Dickson and Company, Sparks, MD, USA)

agar and grown overnight at 36 °C. The MRSA N315 bacterial inoculum was adjusted to approximately 1×10^8 CFU/ml (corresponding to $OD_{540} = 0.5$), then diluted to give 2×10^7 CFU/ml. A 10% porcine mucin (Sigma-Aldrich, St. Louis, MO, USA) suspension was prepared and adjusted to pH 7. Immediately prior to infection, the bacterial inocula were diluted 1:1 with 10% mucin to a final concentration of 1×10^7 CFU/ml in 5% mucin. The mice were then infected IP with 0.5 ml of this inoculum. *In vivo* dosing of compounds in mice was compared with mean or range peak human plasma concentrations of studied β-lactams (165-169).

Antibiotic preparation. Meropenem was obtained from AK Scientific, Inc. (Union City, CA, USA), piperacillin and tazobactam were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA). Linezolid (CAS 165800-03-3) was obtained from AmplaChem (Carmel, IN, USA). Antibiotics were dissolved at a concentration of 16.67 mg/ml in 30% DMSO/30% propylene glycol/40% water. Linezolid was used as positive control and was prepared at 7.5 mg/ml. Vehicle (30% DMSO/30% propylene glycol/40% water) was included as negative control. The dosing formulations were sterilized by passing through 0.2 μm filter prior to injection.

Bacterial isolation from blood. Blood samples were checked for bacterial growth by plating and liquid culture. Whole blood (100 μl, three samples per group) was spread onto Brain-Heart Infusion (BHI) agar plates and incubated at 36 °C overnight. Colonies were counted and three colonies were selected, grown overnight in liquid BHI culture at 36 °C, then mixed 1:1 with 30% LB-glycerol and stored at -80 °C. The remaining three blood samples of each group (50 μl) was added to 5 ml BHI broth and incubated overnight at 36 °C. When growth was noted, cultures were mixed 1:1 with 30% LB-glycerol and stored at -80 °C.

Statistical Analysis

Data for minimal inhibitory concentrations (MICs) are derived from triplicate measurements. Adaptation data are taken from two replicate experiments for each drug combination condition. Data for qRT-PCR expression profiling are derived from three replicate experiments taken from three biological replicates each, with standard error of measurement calculated. Mice were treated in groups of six, and growth determination of bacteria determined via plate and broth culture in triplicate. Fisher's Exact test with Bonferroni correction was used for 8 independent tests (comparing each treatment to vehicle).

A1.4 Results

Synergy between meropenem, piperacillin, and tazobactam in MRSA strains in vitro

Based on its high level of resistance against 23 diverse antibiotics (Table A1.1), *S. aureus* MRSA N315 (170) was selected from a group of fully genome-sequenced MDR strains of MRSA for this study. MRSA N315 contains the staphylococcal chromosome cassette *mec* (SCC*mec*) type II encoding the *mec* methicillin-resistance operon (171), as well as penicillinase plasmid pN315 containing the *bla* β-lactamase operon (172). From a focused combinatorial screen of these 23 antibiotic compounds, including representatives from every major drug class (Table A1.1), we identified the combination of ME/PI/TZ to display highly synergistic, bactericidal activity against MRSA N315 *in vitro*, using the metric of the fractional inhibitory concentration index (FICI), FICI = 0.11 (173, 174) (Table A1.2a). For any number of drugs in combination, a FICI less than 1 indicates synergy, a FICI equal to 1 indicates additivity, and a FICI greater than 1 indicates indifference or antagonism (173, 174). Notably, these three drugs all belong to different sub-classes of the β-lactam drugs, which target the critical transpeptidase enzymes of cell-wall synthesis, though MRSA strains are typically highly resistant to most β-lactams (132). The general resistance to individual β-lactams results from the inability of these drugs to inhibit the

transpeptidase active site of PBP2a, which compensates for β-lactam inhibition of the other transpeptidases in *S. aureus* (132).

ME/PI/TZ exhibits increased synergy against MRSA N315 relative to its three constituent double combinations meropenem/piperacillin (ME/PI), meropenem/tazobactam (ME/TZ) and piperacillin/tazobactam (PI/TZ) at clinically relevant concentrations (Fig. A1.1, Tables A1.2b, c). All three β-lactam compounds were tested for final MIC and FICI using a 3-D checkerboard with twofold dilution series of each compound from 128-to-2 μg/ml, and no-drug. These allowed up to a 64-fold difference in component ratios to be explored for maximal synergy, as well as allowing for isolation of results for each single compound, all constituent double combinations, and the triple combination. Using the 3-D checkerboard, we determined the optimal ratio for ME/PI/TZ to be 1:1:1 for minimal drug input and maximal synergy against MRSA N315. The minimal inhibitory concentrations (MICs) of the three components in the combination against MRSA N315 (2 μg/ml each) are below the clinical susceptibility breakpoints for each of these drugs alone against methicillin-susceptible *S. aureus* (4-8 μg/ml) (175). The constituent double combinations ME/PI and PI/TZ are also synergistic against N315 with FICI = 0.44 and 0.22, respectively, while ME/TZ is less synergistic at 0.67. Based on the Loewe additivity model of synergy, drugs cannot be synergistic with themselves (158). Though the β-lactams all target the cell-wall synthesis pathway, our use of the FICI method (Loewe additivity) confirms the non-additive nature of these interactions. In contrast to the high synergy of ME/PI/TZ seen in MRSA N315, the combination exhibits slightly less than additive activity (FICI = 1.12) in the methicillin-susceptible *S. aureus* (MSSA) reference strain ATCC 29213 (159, 175) (Tables A1.2b, c), and we hypothesize the necessity of PBP2a for synergy to occur.

We propose that the mechanism of synergy observed for ME/PI/TZ results from allosteric triggering of PBP2a by its constituents, akin to that reported for ceftaroline (134, 135). Indeed, we measured that meropenem binds to the allosteric site of PBP2a with a dissociation constant (K_d) of 270 ± 80 μM (equivalent to 104 ± 31 μg/ml). The mean peak plasma concentration in healthy humans after a bolus intravenous (IV) injection of meropenem at the recommended 1 g dose is 112 μg/ml (168). The concentrations of meropenem achieved clinically are above the K_d; thus at these concentrations meropenem binding to the allosteric site of PBP2a would trigger opening of the active site of PBP2a, enabling access to its transpeptidase active site for acylation/inactivation either by another molecule of meropenem or by other β-lactams in the combination (132, 134, 176).

The highly synergistic activity of ME/PI/TZ against MRSA N315 was recapitulated against all of a panel of 72 clinical MRSA isolates with multiple SCC*mec* types represented (Tables A1.3a, b). The MIC of the combination against the clinical isolates ranged from 0.4-33.3 μg/ml for each component, with a mean of 9.7 μg/ml, and an MIC_{50} and MIC_{90} of 3.7 μg/ml and 33.3 μg/ml, respectively (Table A1.4a).

Mechanistic robustness of synergy using alternate carbapenems, penicillins, and β-lactamase inhibitors against MRSA

We determined that the observed synergy is not limited to the antibiotics assayed, but can be generalized to their respective β-lactam classes, by testing MRSA N315 and representative clinical MRSA isolates against other carbapenem/penicillin/β-lactamase inhibitor combinations. We found that treatment of MRSA N315 with imipenem/piperacillin/clavulanate (IM/PI/CV) shows equal or greater synergism to ME/PI/TZ. Meropenem/amoxicillin/tazobactam

(ME/AX/TZ) maintains high synergy in MRSA N315 only (FICI = 0.04), with a clinical MRSA isolate showing less synergy (FICI = 0.55) (Table A1.2b). MICs for components of these substituted triples are all below the mean peak human plasma concentrations of these compounds *in vivo* (165, 166). Similar to ME/PI/TZ, IM/PI/CV shows less-than-additive activity against MSSA ATCC 29213 (FICI = 1.14) (Tables A1.2b, c). These results further support the necessity of the presence of the *mecA* gene product PBP2a with its attendant allosterism for synergy, due to lack of synergy of carbapenem/penicillin/β-lactamase inhibitor combinations in methicillin-susceptible *S. aureus*.

We also tested the effect of replacing the carbapenem component of the combination with either a monobactam or a cephalosporin, two other later-generation β-lactam derivatives. In contrast to ME/PI/TZ, the triple combinations aztreonam/piperacillin/tazobactam (AZ/PI/TZ) and cefepime/piperacillin/tazobactam (CP/PI/TZ) (FICI for both = 0.33) have lower levels of synergy than PI/TZ alone (FICI = 0.22) (Table A1.2b), possibly because aztreonam (a monobactam) has Gram-negative PBP activity (177), while cefepime (a cephalosporin) preferentially targets PBP2 over PBP1 (132).

We confirmed the targets of the constituents of ME/PI/TZ by reducing the expression of PBP1, PBP2, PBP2a or PBP3 using a xylose-inducible antisense-RNA strategy in the MRSA COL strain background (178). When expression levels of PBP2a were attenuated, the strain behaved as a methicillin-susceptible *S. aureus* and was sensitized to all tested β-lactams (Fig. A1.S1a, b, c). When meropenem, piperacillin, and tazobactam were tested against the *pbpA* antisense strain, only meropenem showed larger zones of inhibition under xylose induction, confirming PBP1 as a target of meropenem (Fig. A1.S1d, e). For the *pbp2* antisense strain both meropenem and piperacillin showed increased effectiveness under xylose induction,

demonstrating that they each have some activity against PBP2 (Fig. A1.S1f, g). We did not observe any effect with the *pbp3* antisense strain, consistent with our hypothesis that ME/PI/TZ activity is focused on disrupting PBP1, PBP2, and PBP2a (Fig. A1.S1h, i). The antisense strains in all cases but that of *pbp3* showed sensitization to the triple combination, underscoring the observed synergy.

Lack of adaptation to meropenem/piperacillin/tazobactam over >10 days for MRSA N315

It is obvious that development and spread of resistance can dramatically dampen the effectiveness and longevity of an antimicrobial therapy. We demonstrated that ME/PI/TZ suppresses the evolution of resistance in MRSA using serial passaging in sub-inhibitory antibiotic concentrations of the triple combination and each of its constituents. To more accurately model a clinical treatment *in vitro* and *in vivo*, we applied these drugs at fixed dosages over extended periods as occurs in clinical treatment, not at increasing doses over time. During the 11-day experiment, we observed no evolution of resistance in MRSA N315 to ME/PI/TZ. In contrast, we observed resistance evolution against all double combinations and single constituents within 1-8 days, consistent with prior work (147, 179) (Fig. A1.2). Viable cells were observed in all conditions above the initially determined MIC for the doubles and singles, but not for those conditions at or above the initial MIC for ME/PI/TZ. Increases in growth rates over time were noted in all doubles and singles, while the growth rate of N315 in sub-MIC ME/PI/TZ over time was unchanged throughout the experiment, equivalent to the no-drug control (Fig. A1.2) (147). Also, N315 exposed to the double combination ME/PI showed a threefold increase in MIC after day one, indicating that viable cells were present after day one, but did not grow until further passage and adaptation. Determination of the minimal bactericidal concentration (MBC) confirmed that the triple combination ME/PI/TZ is bactericidal against MRSA N315

(Table A1.4b). Together, these results demonstrate the suppression of emergence of new resistance against ME/PI/TZ in MRSA N315.

Reciprocal collateral sensitivities of components of these combinations underlie suppression of adaptation

To determine whether collateral sensitivity was a factor in the suppression of adaptation of ME/PI/TZ, we analyzed the effects of prior exposure of MRSA N315 to a range of β-lactams on susceptibility to the other components (Fig. A1.3 and Fig. A1.S2). We observed that there was strong reciprocal collateral sensitivity between meropenem and piperacillin, and between piperacillin and ME/TZ, while PI/TZ sensitized MRSA N315 to meropenem, but not reciprocally. Collateral sensitivity to piperacillin was also conferred by prior exposure to tazobactam, but not vice-versa. Interestingly, no collateral sensitivity was found to tazobactam after exposure to any other single or double compounds. Collateral sensitivity and resistance profiles of amoxicillin and piperacillin are nearly identical, with adaptation to meropenem also sensitizing MRSA N315 to amoxicillin (Fig. A1.3 and Fig. A1.S2). Piperacillin also showed collateral sensitization to imipenem, an even more potent carbapenem against MRSA N315. However, none of the cephalosporins tested for collateral sensitivity by the carbapenem/penicillin/β-lactamase inhibitor combinations or constituents resulted in sensitivity, but rather increased resistance or indifference was noted. These results confirm that the observed suppression of resistance by collateral sensitivity is specific to the constituent drug classes of ME/PI/TZ.

Adapted MRSA N315 undergoes large-scale genomic alterations

We used whole-genome sequencing to investigate the genomic basis of the sensitivity and resistance phenotypes of wild-type and adapted MRSA N315 strains. We found no mutations in PBP or β-lactamase genes within any of the adapted MRSA N315 isolates. However, absence of read coverage identified that the penicillinase plasmid pN315 was lost in isolates adapted to tazobactam-only (100 μg/ml) and ME/TZ (11.1 μg/ml each) (Fig. A1.4a). This plasmid loss occurred much more rapidly than with previously reported techniques for curing plasmids from MRSA, such as high heat and SDS treatment (180). In PI/TZ adapted isolates, we observed that approximately 400 kb of the MRSA N315 chromosome (GenBank ID: BA000018.3) was duplicated after analysis of read coverage depth, from approximate genomic positions 2,100,000 to 2,550,000 bp. Interestingly, this interval contains several putative and confirmed genes involved in cell-wall synthesis, including *ddlA* D-Ala-D-Ala ligase (Fig. A1.S3).

The loss of pN315 in MRSA N315 correlates with increased sensitivity to piperacillin and amoxicillin, both penicillins that should be sensitive to the *blaZ* (PC1) class A β-lactamase encoded on the plasmid. However, the loss of pN315 also results in increased resistance to tazobactam-only and ME/PI/TZ (Fig. A1.3, Fig. A1.S2, Table A1.5a). One possible link between the presence of pN315 and ME/PI/TZ activity is the known regulatory crosstalk between MecI and BlaI repressors and their shared *mec* operon target (181-183). To test the effect of the loss of pN315 on expression of genes known to be important for ME/PI/TZ activity, we performed qRT-PCR analysis of the adapted and wild-type MRSA N315 strains (Fig. A1.4b). We determined that expression of the *blaZ* β-lactamase in the pN315 plasmid within wild-type MRSA N315 is constitutive, but in clones adapted to tazobactam we saw no expression of *blaZ*, consistent with loss of pN315 in these clones. We also found that expression of *mecA* is constitutive in the *blaZ*-null MRSA N315 isolate that was adapted to tazobactam at 100 μg/ml, consistent with

disregulation of the *mec* operon via loss of pN315 and the *bla* operon. Finally, we found tazobactam to be a strong inducer of *mecA* in wild-type MRSA N315, at levels similar to the constitutive expression of *mecA* seen in the *blaZ*-null condition.

Synergy of ME/PI/TZ when MRSA N315 has evolved resistance to constituents

We then examined the role that resistance to components of ME/PI/TZ has on its effectiveness against MRSA (Table A1.5a). Previous exposure of MRSA N315 to piperacillin at either 33.3 or 100 μg/ml showed subsequent sensitization of the strain to ME/PI/TZ, from 3.7 to 1.2 μg/ml for each component. However, prior exposure of MRSA N315 to ME/TZ (11.1 μg/ml each) or meropenem-only (33.3 μg/ml) showed a nine-fold increase in levels of resistance to ME/PI/TZ (increasing from 3.7 to 33.3 μg/ml for each component). Exposure to tazobactam-only gave intermediate gains in resistance to ME/PI/TZ up to day 7 (11.1 μg/ml each), and higher resistance at day 11 (33.3 μg/ml each). Prior exposure to ME/PI or PI/TZ generated only a threefold increase in MIC (from 3.7 to 11.1 μg/ml) over the 11 days.

Despite the elevated MICs to ME/PI/TZ in the isolates adapted to the component drugs, the triple-drug combination still maintained synergy in all adapted isolates (Table A1.5b). This is consistent with synergistic drug activity within the range of ME/PI/TZ MICs observed for the 72 clinical MRSA isolates (Table A1.4), relative to their single-drug MICs. These results show that even when genomic changes enabling sub-component resistance can be selected, the overall synergistic activity of the triple-drug combination is maintained. In contrast to recent work with a non-pathogenic *E. coli* strain (158), we observed no change in the overall drug interaction profile of ME/PI/TZ regarding synergy with increased resistance to any component drug.

***ME/PI/TZ is as effective as linezolid against MRSA* in vivo**

Next we tested if ME/PI/TZ or its constituents can be effective in treating MRSA infections *in vivo* using a neutropenic mouse model of peritonitis. Blood taken at 11 h post-infection from mice that were treated with either ME/PI/TZ, ME/PI (67 mg/kg each) or linezolid (30 mg/kg) (184) yielded zero plated colonies and no growth in liquid cultures, indicating clearance of infection (Fig. A1.5, Fig. A1.S5). All mice (n = 6/group) from each of these treatments survived for six days post-infection (total duration of the mouse study). The activity of ME/PI/TZ and ME/PI was similar to linezolid monotherapy based on clearance of MRSA infection and complete survival of all treated mice compared to vehicle ($p = 0.02$, Fisher's exact test).

In contrast to the complete rescue of the infected mice by ME/PI/TZ, ME/PI, or linezolid, several mice treated with ME/TZ, PI/TZ, or meropenem-alone, and all mice treated singly with piperacillin or tazobactam succumbed to the infection, most within 48 h (Fig. A1.5). Treatment with these other drug regimens was not significantly different than treatment with vehicle-only ($p > 0.05$, Fisher's exact test) (Table A1.6a), where all mice also succumbed to the infection within 48 h.

We tested MRSA N315 cultures from blood drawn from mice treated with meropenem, piperacillin, or vehicle for their *in vitro* MICs against ME/PI/TZ and its constituent single drugs to determine whether adaptation occurred during passage *in vivo*. All four tested isolates of MRSA N315 had identical MICs for the triple ME/PI/TZ and all constituent drugs, and thus identical synergy (Table A1.6b). These data suggest no adaptation occurred within these strains to overcome the triple ME/PI/TZ tested within the 11-h passage *in vivo*.

A1.5 Discussion

We have shown that triple antibacterial combinations containing carbapenems, penicillins, and β-lactamase inhibitors target multiple nodes in the same cellular system (cell-wall synthesis) and are highly synergistic and bactericidal against diverse MRSA strains *in vitro*, at clinically achievable concentrations. This contrasts with recent work showing collateral sensitivity and synergy to arise from combinations of drug classes working against orthogonal cellular targets in non-pathogenic lab strains only (145, 158). Because carbapenems and other drugs at high concentration could have toxic effects, reduced per-drug dosages via synergy mitigate potential toxicities (185). Our 3-D checkerboard testing confirmed the optimal input concentrations for ME/PI/TZ to be given in a 1:1:1 ratio (2 μg/ml each) against MRSA N315, which is below the susceptibility breakpoints for these compounds against methicillin-susceptible *S. aureus*, and is an 8-to-64-fold reduction in input concentrations for these formerly inactive drugs against this highly resistant MRSA strain. Our mechanistic analyses support our hypothesis that targeting of PBP1 by meropenem, targeting of PBP2 by piperacillin, protection of piperacillin by tazobactam from β-lactamase cleavage, and allosteric opening of the active site of PBP2a by meropenem for inhibition by another molecule of antibiotic in the combination, result in synergy by simultaneously perturbing multiple components of the MRSA cell-wall synthesis system (Fig. A1.S4).

We have also preliminarily shown that this combination has activity in a highly lethal neutropenic MRSA *in vivo* model, demonstrating that this triple combination of clinically approved β-lactams can clear infection similar to a substantially more expensive monotherapy like linezolid. The plasma levels of meropenem observed in mice correlate well with plasma drug levels in healthy humans (186), and meropenem would attain the K_d at these clinically

achievable concentrations to trigger allostery for opening of the active site of PBP2a, providing accessibility for inhibition by meropenem and other β-lactams in the combination (133, 134).

Notably, the double combination ME/PI cleared the MRSA N315 infection *in vivo* similarly to ME/PI/TZ and linezolid within 11 h. *In vitro* we observed high synergy scores and reciprocal collateral sensitivity for this combination, similar to what was seen for ME/PI/TZ, but ME/PI did not suppress evolution of resistance to the same extent that ME/PI/TZ did. This property may not have been relevant to this aggressive infection model, but may be important for longer treatment times seen in human infections with MRSA. ME/PI/TZ is also likely to be effective at lower total concentrations than ME/PI because of its higher synergy. Longer exposure of the N315 strain to the tazobactam component of ME/PI/TZ *in vivo* may also promote ejection of pN315 plasmid with concomitant sensitization to the penicillin component, in line with the *in vitro* results for collateral sensitivity and suppression of adaptation. Indeed, to more adequately address this question, potential longer-term *in vivo* resistance evolution would need to be tested under sub-lethal concentrations of the drugs in important follow-up mouse experiments.

Our robust mechanistic *in vitro* results and preliminary *in vivo* results for ME/PI/TZ activity suggest this combination may be made immediately available for use in the clinic, since it includes currently FDA-approved drugs, which had met their obsolescence as monotherapies against MRSA decades ago. However, further mechanistic features of the combination that were shown *in vitro* (synergy, resistance suppression over longer periods of dosing, collateral sensitivity, etc.) will require substantially more *in vivo* testing to support the promising but preliminary activity observed in our highly aggressive neutropenic mouse model.

We note that high resistance to meropenem or tazobactam slightly reduces the effectiveness of ME/PI/TZ, while maintaining its synergy, and our resistance evolution analysis

cannot account for resistance genes acquired horizontally that could break the relationship between meropenem, piperacillin, and tazobactam. Despite these caveats, we believe the ME/PI/TZ combination is an immediately viable anti-MRSA therapeutic, and endorse further mechanistic exploration into the putative superior efficacy of high-order antibiotic combinations that are both synergistic and encoded by collaterally sensitive constituents. Having similar activity to linezolid against MRSA *in vivo*, the potential efficacy of ME/PI/TZ reopens broad prospects for the clinical use of β-lactams against the staphylococci. It also suggests that this line of research into repurposing existing antibiotics in carefully designed synergistic combinations would address immediate clinical needs, as these agents are already approved for human use. Emergence of resistance to any antibiotic or any antibiotic combination is inevitable. Yet, as evidenced in our study, combinations composed of key drug-drug interaction features may be a tool in mitigating the emergence of antibiotic resistance by preserving the usefulness of existing agents available to us in our pharmacological armamentarium.

A1.6 Figures

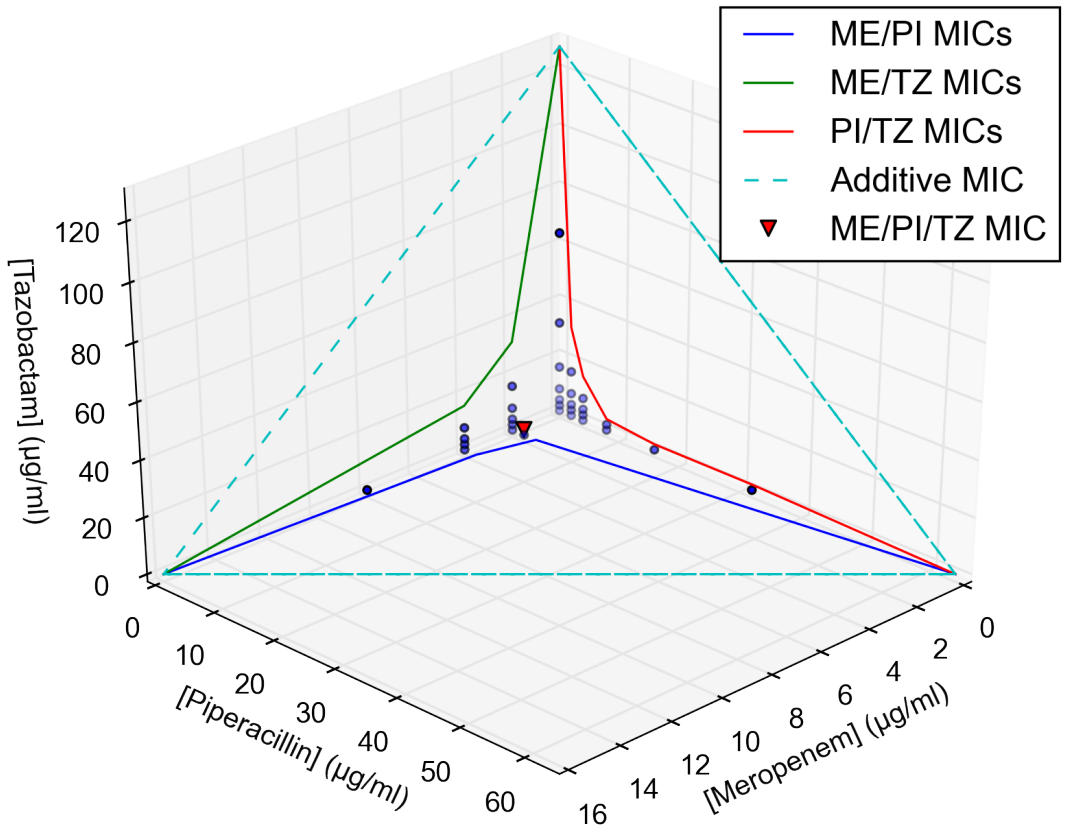


Figure A1.1

3D-Checkerboard synergy determination showing isoboles of minimal inhibitory concentrations (MIC) and *in vitro* growth in single-, double-, or triple-drug conditions for ME/PI/TZ. Colored lines/isoboles within each panel indicate MICs of two drugs in combination. Dashed lines indicate theoretical concentrations of additive interactions. Points indicate top sub-inhibitory concentrations of meropenem (ME), piperacillin (PI) and tazobactam (TZ) for each tested condition. The red triangle indicates the MIC of all three drugs in combination (Each at 2 µg/ml).

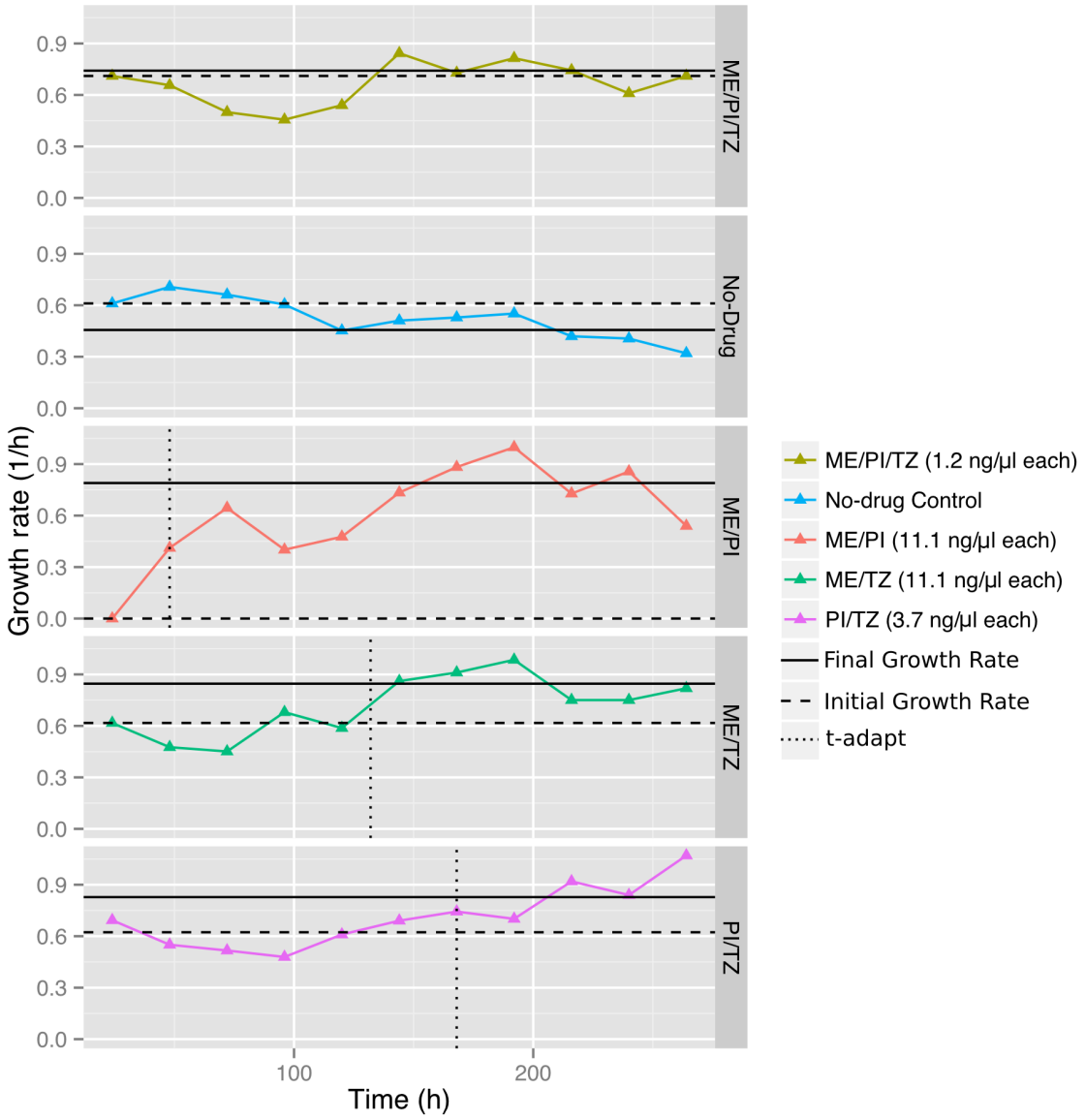


Figure A1.2

Change in growth rates over time of MRSA N315 when challenged with antibacterial combinations. Growth rates of MRSA N315 over an 11-day period were computed for each antibacterial combination tested at one threefold dilution below MIC. The differences in growth rate (Δr) between day one (Initial Growth Rate) and the averaged rate of the last six days of the assay (Final Growth Rate) were calculated. MRSA N315 in conditions whose change in growth

rate Δr >0.2 were considered to be adapted. The adaptation time parameter *t-adapt* was calculated as the time at which change in growth rate was half-maximal. Adaptation rate, $\alpha = (\Delta r/2)/t\text{-}adapt$ ($1/h^2$), was computed for strains meeting this criterion. Results are from two replicate experiments. Adaptation rate for ME/PI: $\alpha = 8.23\times10^{-3}$ h^{-2}; ME/TZ: $\alpha = 8.68\times10^{-4}$ h^{-2}; PI/TZ: $\alpha = 4.32\times10^{-3}$ h^{-2}. Only ME/PI/TZ at one threefold dilution below MIC (1.2 µg/ml each) and No-drug Control displayed lack of increase in growth rate and were non-adapted.

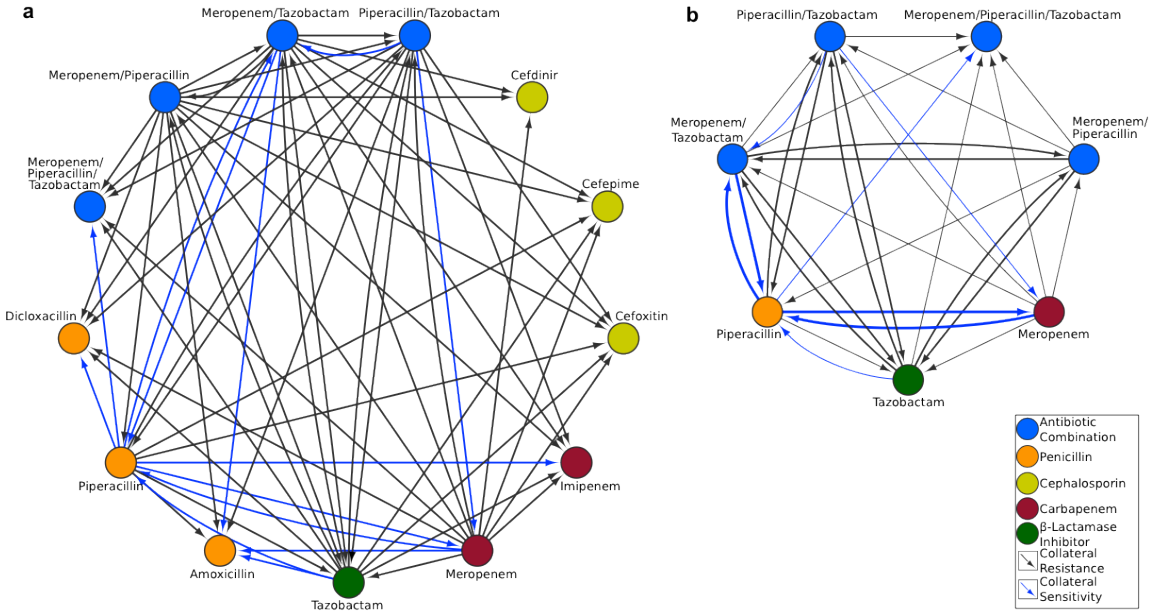


Figure A1.3

Collateral sensitivities underlie suppression of adaptation to antibacterial combinations in MRSA N315. a, MRSA N315 interaction network of collateral sensitivities and resistance between ME/PI/TZ, its single and double constituents, and other β-lactam compounds of various sub-classes (cephalosporins, penicillins, carbapenems, and β-lactamase inhibitors). Node colors indicate sub-classes of β-lactams, β-lactamase inhibitors, or combinations. Blue arrows indicate

collateral sensitivities. Black lines indicate collateral resistance. For example, adaptation to piperacillin sensitizes MRSA N315 to meropenem and imipenem. Cephalosporins were not collaterally sensitive to any of the compounds we tested. Where pairs were not tested or no collateral effects were seen, no connecting arrows are shown. **b**, MRSA N315 interaction network of collateral sensitivities and resistance between ME/PI/TZ and its single and double constituents only. Bold blue arrows indicate reciprocal collateral sensitivities between two nodes, e.g., piperacillin and meropenem/tazobactam.

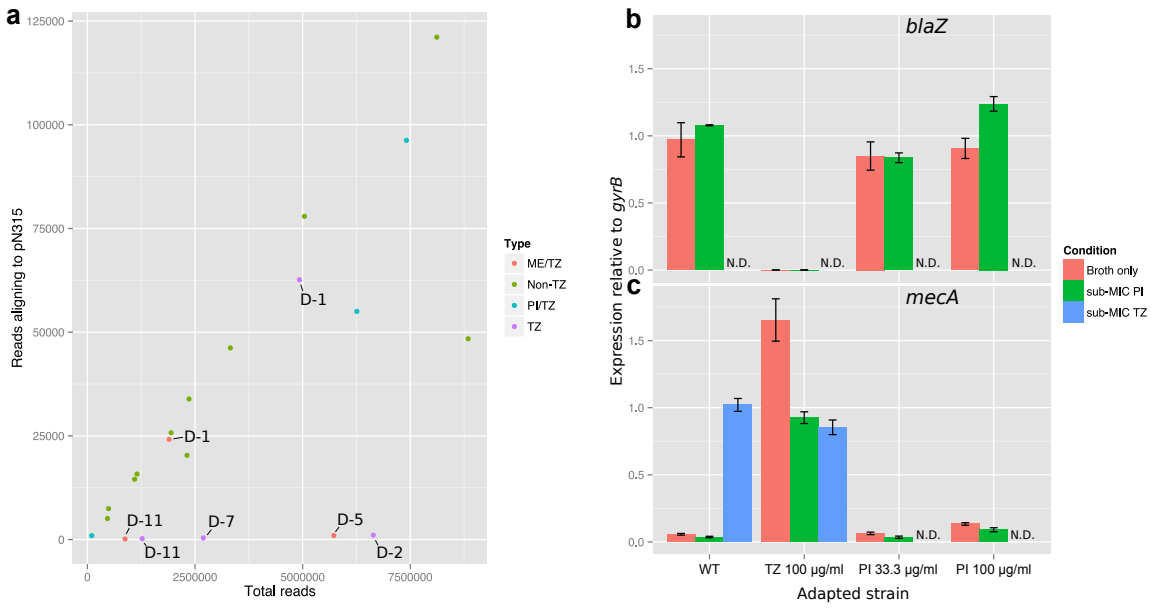


Figure A1.4

Genomic evidence for mechanisms of synergy and collateral sensitivity. a, Adaptation of MRSA N315 to meropenem/tazobactam or tazobactam alone destabilizes plasmid pN315. Read coverage aligning to pN315 in MRSA N315 adapted to drug combinations containing tazobactam (TZ) or not containing tazobactam (non-TZ), versus total reads per sample. Days of adaptation under the given conditions are indicated, e.g., D-2 indicates isolate was sequenced after two days of adaptation. b, c, qRT-PCR confirms disregulation of the *bla* and *mec* operons

as causative mechanisms of some collateral sensitivities in MRSA N315. Expression of *blaZ* or *mecA* shown relative to *gyrB* in wild-type MRSA N315 or adapted strains (N315 adapted to TZ 100 μg/ml, and PI 33.3 or 100 μg/ml), subsequently grown in broth-only or broth + sub-MIC PI or TZ. N.D. = Not determined. "-" indicates no expression. Loss of *blaZ* expression in MRSA N315 adapted to TZ confirms loss of *blaZ* and the *bla* operon, and is consistent with disregulation of *mecA* expression. Data are from three replicate experiments. Error bars indicate standard error of measurement.

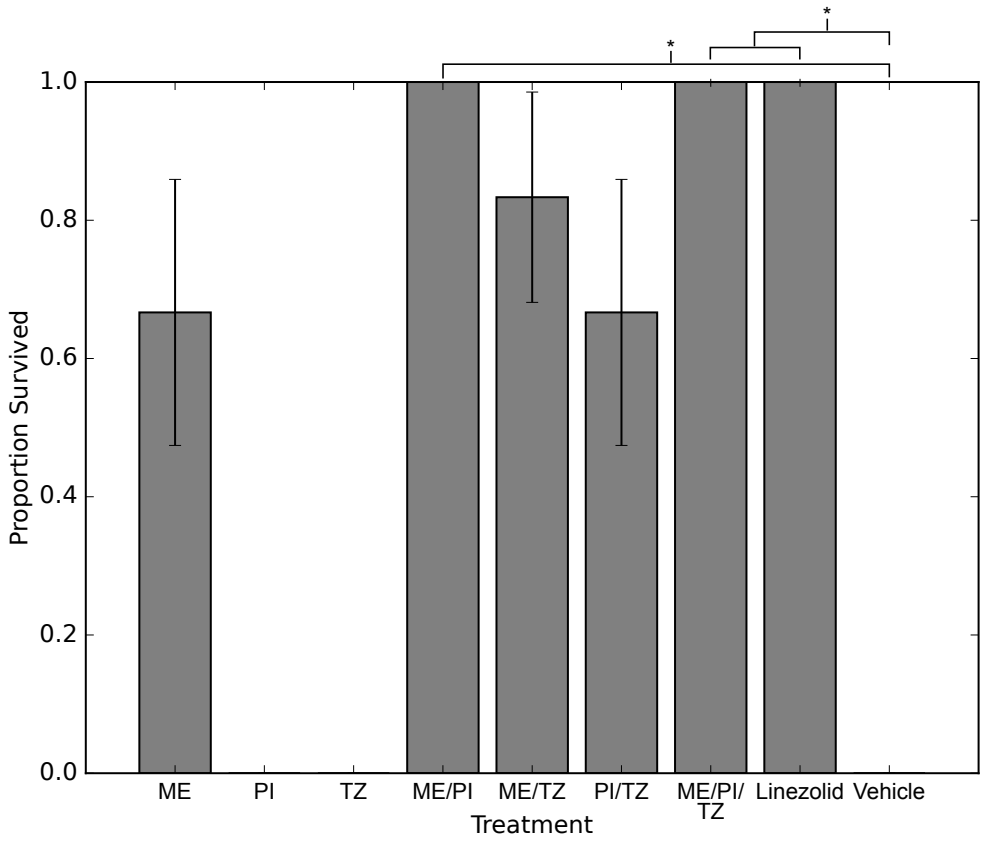


Figure A1.5

Efficacy of ME/PI/TZ treatment in a neutropenic mouse peritonitis model of MRSA N315. Proportional survival of mice (n = 6) from each drug treatment is shown. Treatment with

ME/PI/TZ, ME/PI, and linezolid are significantly different than vehicle (*p = 0.02). Error bars indicate standard error of proportions of survivors per condition tested.

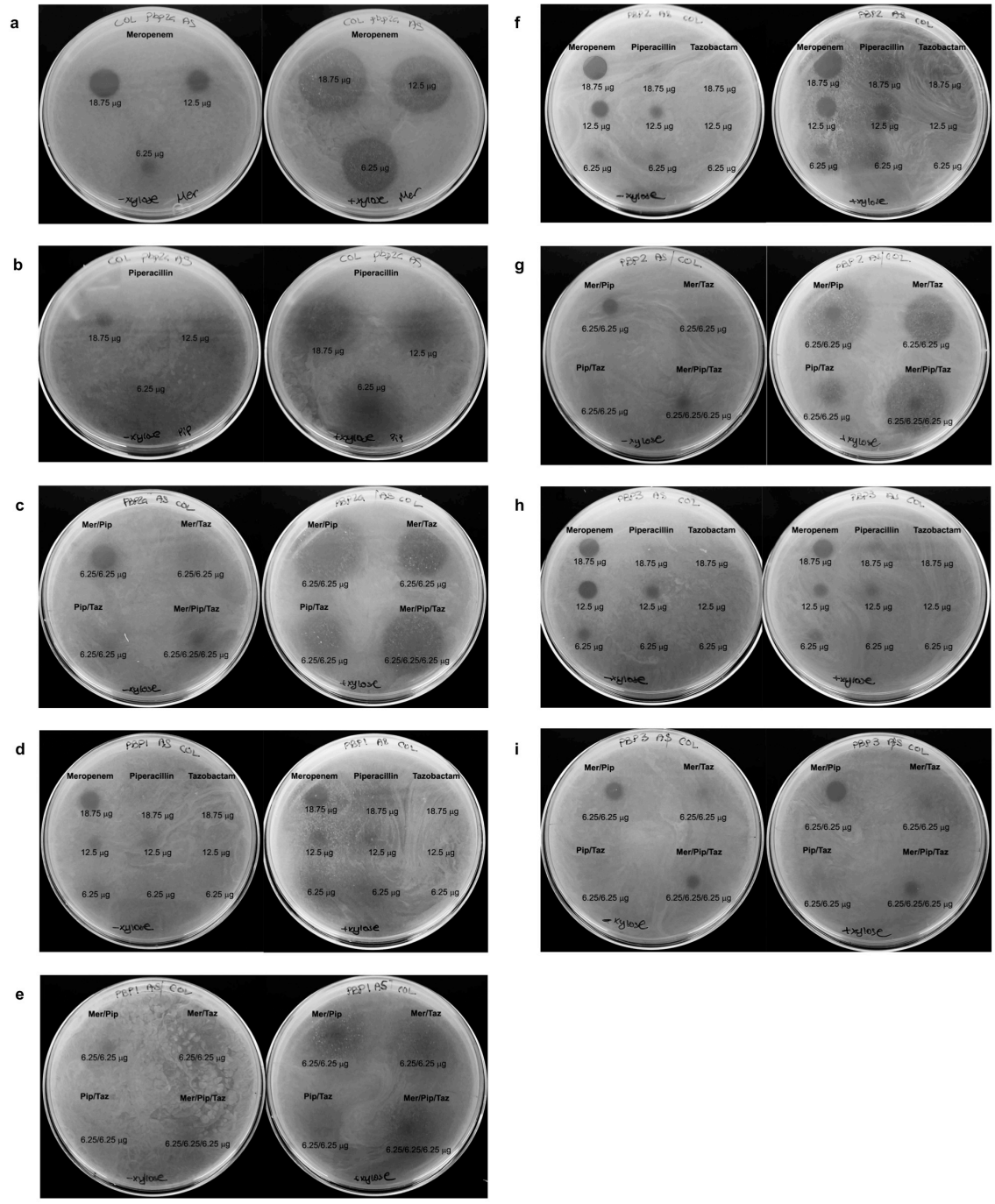


Figure A1.S1

PBP xylose induction in MRSA COL antisense (AS) strains. a, b, c, *pbp2a/mecA* antisense (AS) strain. Targeted repression of PBP2a showed increased susceptibility for meropenem when under xylose induction. Increased susceptibility was also observed for piperacillin. Increased susceptibility was observed for all combinations. d, e, *pbpA* antisense (AS) strain. Targeted repression of PBP1 showed increased susceptibility to meropenem and piperacillin. Increased susceptibility was observed for the ME/PI, ME/TZ, and ME/PI/TZ combinations. f, g, *pbp2* antisense (AS) strain. Targeted repression of PBP2 showed increased susceptibility to meropenem and piperacillin. Increased susceptibility was observed for all combinations. h, i, *pbp3* antisense (AS) strain. Targeted repression of PBP3 showed no increase in susceptibility to any of the single drugs. A slight increase in susceptibility was observed for the ME/PI combination; no change in susceptibility was observed for any of the other combinations.

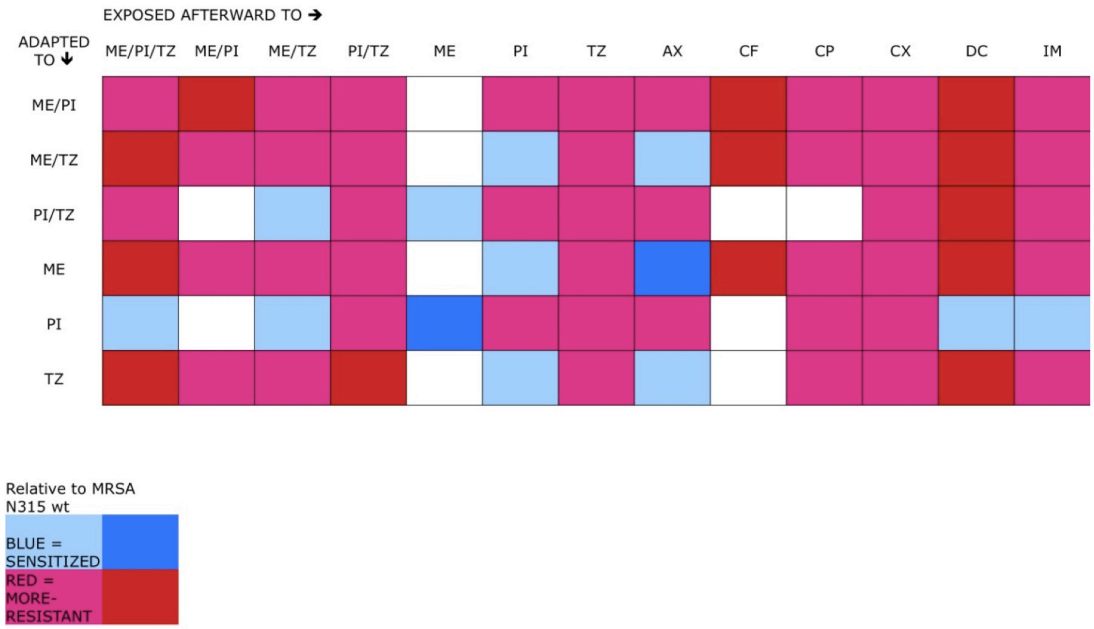


Figure A1.S2

Collateral sensitivities underlie suppression of adaptation to β-lactam combinations in MRSA N315. Blue shades indicate collateral sensitization of strain to single drugs and combinations, after prior adaptation to single and double drug combinations. Red shades indicate collateral resistance. Light shading = change of one MIC dilution. Dark shading = change of two or more MIC dilutions. For example, adaptation to piperacillin yields collateral sensitivity to meropenem, and vice versa. ME = meropenem, PI = piperacillin, TZ = tazobactam, AX = amoxicillin, CF = cefdinir, CP = cefepime, CX = cefoxitin, DC = dicloxacillin, IM = imipenem.

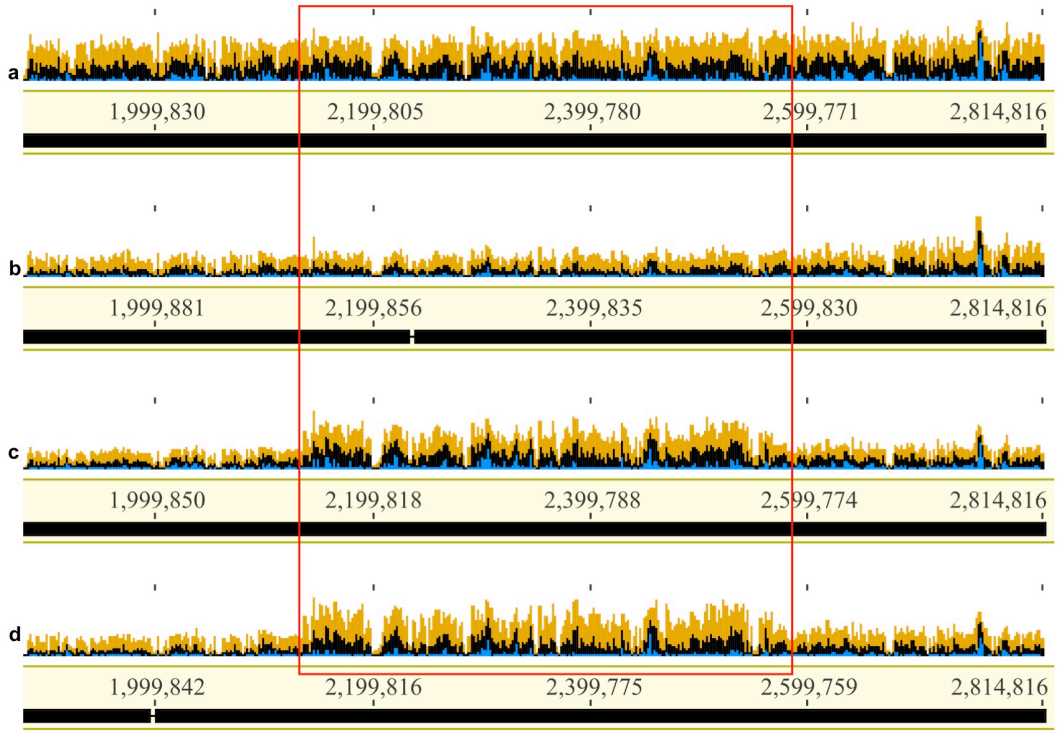


Figure A1.S3

Genomic duplication in MRSA N315 adapted to piperacillin/tazobactam. Histogram showing the total read coverage across the genome of N315 adapted to a, meropenem/tazobactam for five days, b, tazobactam alone for two days, c, piperacillin/tazobactam for six days, and d,

piperacillin/tazobactam for 11 days. Average per-base read coverage across the entire genome and only in the region indicated by the red box are, respectively: a) 116.6 reads/bp and 126.6 reads/bp; b) 124.5 reads/bp and 128.9 reads/bp; c) 157.8 reads/bp and 302.2 reads/bp; and d) 120.1 reads/bp and 230.6 reads/bp. Clones in a and b were chosen to be representative of all non-piperacillin/tazobactam adaptations.

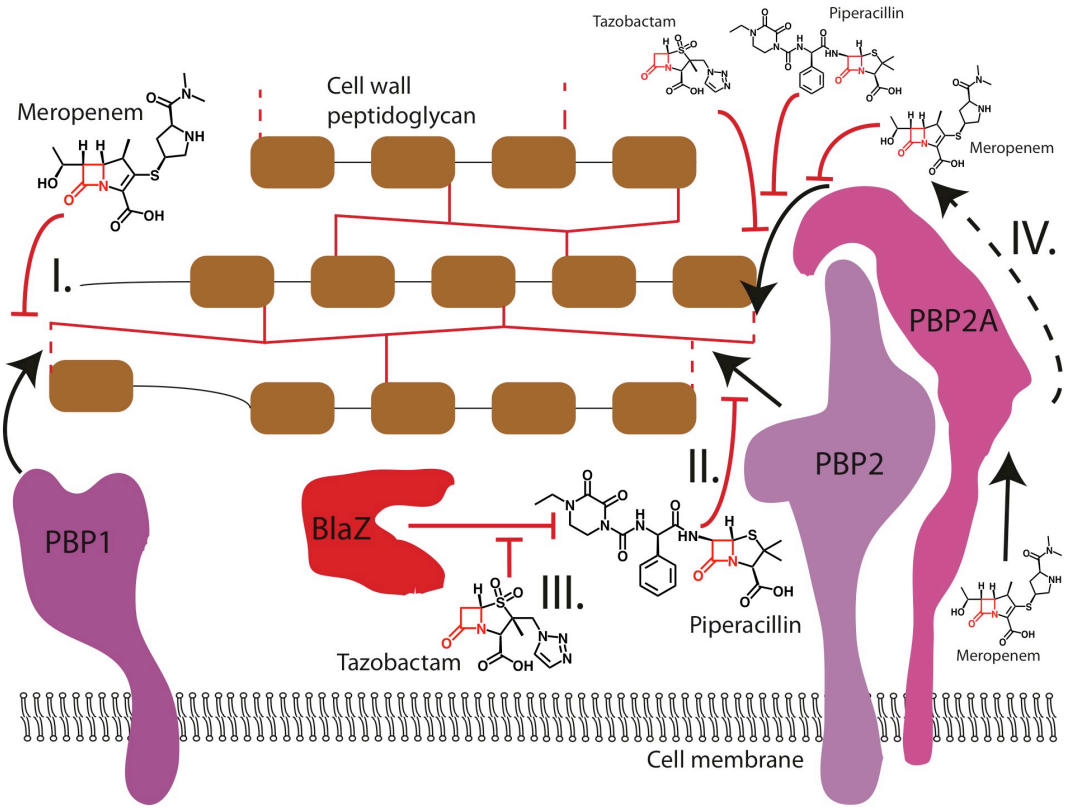


Figure A1.S4

Proposed mechanism of synergy of meropenem/piperacillin/tazobactam (ME/PI/TZ) against MRSA. Our data support the proposed synergistic mode of action against cell-wall synthesis in MRSA involving: I.) suppression of transpeptidation by PBP1 at the division septum by carbapenems, II.) suppression of transpeptidation by PBP2 by penams (penicillins), III.)

suppression of β-lactamase activity against penams by β-lactamase inhibitors, and IV.) allosteric opening of the active site of PBP2a by meropenem, allowing inhibition by meropenem or by other β-lactams.

A1.7 Tables

Compound	Target mechanism in bacteria	Antibiotic Class	MIC in MRSA N315 (µg/ml)
Sulfamethoxazole	folic acid pathway	Sulfonamide	100
Trimethoprim	folic acid pathway	Pyrimidine derivative	6.2
Levofloxacin	DNA synthesis	Fluoroquinolone	0.4
Bleomycin	DNA synthesis	Glycopeptide	>500
Gemfibrozil	lipid synthesis	*Fibrate (hyperlipidemia agent)	>200
Sulfometuron	amino acid biosynthesis	*Broad-spectrum urea herbicide	>200
Disulfiram	osmotic stress response	*Thiuram disulfide (anti-alcohol therapeutic)	11.1
Tigecycline	protein synthesis	Tetracycline	0.4
Mupirocin	protein synthesis	Pseudomonic acid	0.4
Linezolid	protein synthesis	Oxazolidinone	3.7
Azithromycin	protein synthesis	Macrolide	>200
Clindamycin	protein synthesis	Lincosamide	>500
Chloramphenicol	protein synthesis	Amphenicol	11.1
Tobramycin	protein synthesis	Aminoglycoside	>500
Rifampin	transcription	Rifamycin	0.4
Vancomycin	cell wall synthesis	Glycopeptide	0.4
Piperacillin	cell wall synthesis	β-lactam/Penicillin (Penam)/Broad-spectrum	64
Aztreonam	cell wall synthesis	β-lactam/Monobactam/Gram-negative specific	>500
Cefepime	cell wall synthesis	β-lactam/Cephalosporin 4th generation (Cephem)/Broad-spectrum	100
Meropenem	cell wall synthesis	β-lactam/Carbapenem/Ultra-broad-spectrum	16
Tazobactam	cell wall synthesis	β-lactamase inhibitor (Penam)	128
D-Cycloserine	cell wall synthesis	Analogue of the amino acid D-alanine	56
Colistin	cell membrane lysis	Polymyxin	500

Table A1.1

23 antibacterial compounds used to formulate combinations in this study. Compounds are grouped by target mechanism of action. *Compound not formally classified as an antibiotic drug, but has known antibacterial properties.

a

FICI	Interpretation
≤0.5	Synergy
>0.5 to <1.0	Partial synergy
1.0	Additivity
>1.0 to <4.0	Indifference
≥4.0	Antagonism

b

Combination in strain	ME/PI/TZ	CP/PI/TZ	AZ/PI/TZ	ME/AX/TZ	ME/AX/CV	IM/PI/CV	ME/PI	ME/TZ	PI/TZ	IM/PI	IM/CV	PI/CV
MRSA N315 SCC*mec* type II	0.11	0.33	0.33	0.04	0.41	0.06	0.44	0.67	0.22	0.15	0.67	0.44
MRSA #181 SCC*mec* type II	0.28	ND	ND	0.55	ND	0.11	ND	ND	ND	ND	ND	ND
MSSA ATCC 29213	1.12	ND	ND	ND	ND	1.14	2.97	8.61	0.36	1.11	1.04	0.43

c

Combination in strain	ME/PI/TZ	CP/PI/TZ	AZ/PI/TZ	ME/AX/TZ	ME/AX/CV	IM/PI/CV	ME/PI	ME/TZ	PI/TZ	IM/PI	IM/CV	PI/CV
MRSA N315 SCC*mec* type II	2 each	11.1 each	11.1 each	0.4 each	3.7 each	0.12/1.2/1.2	2/4	8/2	16/2	0.37/3.7	1.11/11.1	11.1 each
MRSA #181 SCC*mec* type II	11.1 each	ND	ND	11.1 each	ND	0.37/3.7/3.7	ND	ND	ND	ND	ND	ND
MSSA ATCC 29213	0.27 each	ND	ND	ND	ND	0.04/0.4/0.4	0.4 each	1.2 each	1.2 each	0.04/0.4	0.04/0.4	1.2 each

Table A1.2

Fractional Inhibitory Concentration Index (FICI) profiling of combinations. a, Interpretive criteria for FICI scoring. b, FICI profiles of various triple combinations of carbapenems/penicillins/β-lactamase inhibitors against MRSA and MSSA strains. c, MIC profiles of same combinations (μg/ml). Constituent double combinations are shown for comparison.

a

Clinical MRSA isolate (SCC*mec* type)	FICI score
4	0.22
7	0.22
13	0.15
15	0.5
22	0.22
25	0.22
27	0.67
31	0.15
35 (II)	0.17
37	0.15
39 (II)	0.67
41 (II)	0.09
45	0.22
48	0.15
53	0.22
59	0.07
64 (II)	0.44
66	0.22
70	0.67
72	0.15
73 (IV)	0.34
74	0.34
75	0.34
77 (II)	0.67
85	0.5
89	0.07
90	0.22
95	0.22
99	0.44
101	0.22
103	0.37
104 (II)	0.15
109	0.17
118	0.07
121	0.22
122	0.22

b

Clinical MRSA isolate (SCC*mec* type) (Continued)	FICI score (Continued)
124	0.28
131 (II)	0.37
132	0.22
140	0.22
144	0.15
146	0.22
150	0.67
152	0.22
155	0.39
161	0.37
163	0.37
164	0.17
165	0.15
167	0.22
168	0.22
169	0.17
171	0.17
172	0.67
175	0.15
177	0.22
181 (II)	0.28
182	0.5
189	0.15
190	0.34
193 (II)	0.34
194	0.15
195	0.15
197	0.15
200	0.15
201	0.44
204	0.39
205	0.22
206	0.22
213	0.44
217	0.15
219	0.15

Table A1.3

Compiled FICI data for ME/PI/TZ against MRSA N315 and 72 clinical MRSA isolates. a, b, 72 clinical MRSA isolates (with SCC*mec* type, if known) and FICI scores for ME/PI/TZ against 72 clinical MRSA isolates.

a

MIC of ME/PI/TZ components (µg/ml)	# of MRSA isolates	% of total
33.3	9	12.3
11.1	27	36.9
3.7	27	36.9
1.2	8	10.9
0.4	2	2.7
Total	73	-

b

Plate Concentrations	Colonies Plate A	Colonies Plate B
ME/PI/TZ 2/2/2	Punctate lawn	Too Many to Count
ME/PI/TZ 4/4/4*	40	0
ME/PI/TZ 8/8/8	2	0
ME/PI/TZ 16/16/16	0	0
ME/PI/TZ 32/32/32	8	0

*MIC = 4/4/4 µg/ml

Table A1.4

Compiled MIC and MBC data for ME/PI/TZ against MRSA isolates. a, Distribution of MIC resistance profiles of studied MRSA isolates against ME/PI/TZ. b, Confirmation of minimum bactericidal concentration (MBC) for ME/PI/TZ in MRSA N315.

a

Passage day /Adaptation conditions	ME/PI 11.1 µg/ml each	ME/TZ 11.1 µg/ml each	PI/TZ 3.7 µg/ml each	Meropenem 33.3 µg/ml	Piperacillin 100 µg/ml	Piperacillin 33.3µg/ml	Tazobactam 100 µg/ml
1	no change	11.1/11.1/11.1	1.2/1.2/1.2	11.1/11.1/11.1	no change	1.2/1.2/1.2	11.1/11.1/11.1
2	no change	no change	no change	no change	3.7/3.7/3.7	3.7/3.7/3.7	11.1/11.1/11.1
3	no change	no change	no change	no change	no change	no change	no change
4	no change	no change	no change	no change	no change	no change	no change
5	no change	33.3/33.3/33.3	no change	no change	no change	no change	no change
6	no change	no change	11.1/11.1/11.1	no change	no change	no change	no change
7	11.1/11.1/11.1	no change	no change	33.3/33.3/33.3	no change	no change	11.1/11.1/11.1
8	no change	no change	no change	no change	1.2/1.2/1.2	1.2/1.2/1.2	no change
9	no change	no change	no change	no change	no change	no change	no change
10	no change	no change	no change	no change	no change	no change	no change
11	11.1/11.1/11.1	33.3/33.3/33.3	11.1/11.1/11.1	33.3/33.3/33.3	1.2/1.2/1.2	1.2/1.2/1.2	33.3/33.3/33.3

b

MRSA N315 isolate adapted to:	ME/PI	ME/TZ	PI/TZ	ME	PI	TZ
$FICI_{ME/PI/TZ}$	0.22	0.83	0.22	0.83	0.05	0.83

Table A1.5

Change in ME/PI/TZ resistance phenotype of MRSA N315 over 11 days after repeated exposure to constituents of ME/PI/TZ. a, Isolates were selected on days when an increase in MIC or growth rate was noted. Antibacterial concentrations listed (in µg/ml) show the adaptation conditions for MRSA N315. Post-adaptation MICs to each component of ME/PI/TZ are shown

in selected isolates versus passage day. b, FICI of MRSA N315 against ME/PI/TZ after adaptation to components *in vitro*.

a

Drug condition tested	p-value versus vehicle	After multiple hypothesis correction (Bonferroni)
ME	0.606	4.848 (1)
PI	1	8 (1)
TZ	1	8 (1)
ME/PI	0.0022	0.0176
ME/TZ	0.0152	0.1216
PI/TZ	0.606	4.848 (1)
ME/PI/TZ	0.0022	0.0176
Linezolid	0.0022	0.0176

b

Colonies from mice given: / MIC (μg/ml) for:	ME	PI	Vehicle	Wild-type N315
ME/PI/TZ	3.7/3.7/3.7	3.7/3.7/3.7	3.7/3.7/3.7	3.7/3.7/3.7
ME	33.3	33.3	33.3	33.3
PI	33.3	33.3	33.3	33.3
TZ	100	100	100	100
FICI	0.26	0.26	0.26	0.26

Table A1.6

a, Statistics of *in vivo* treatments with β-lactams. b, *in vitro* MICs and FICI scores for MRSA N315 after passage *in vivo* under indicated drug conditions.