Genomic Epidemiological Insights Into MRSA Transmission and Adaptation in an Urban Jail and the Surrounding Community

by

Stephanie N. Thiede

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in The University of Michigan
2021

Doctoral Committee:

Assistant Professor Evan Snitkin, Chair
Associate Professor Michael Bachman
Professor Mary O'Riordan
Assistant Professor Krishna Rao
Assistant Professor Jon Zelner

Stephanie N. Thiede

thiede@umich.edu

ORCID: 0000-0003-0173-4324

Dedication

To my Grandma Ann Thiede

1930 - 2013

Acknowledgements

My love for research was borne from the mentorship of Drs. Sherry Voytik-Harbin and the late Ann Rundell, who gave me my first research opportunity as an undergraduate student at Purdue University and recognized me as a scientist before I ever realized I was one. Sherry encouraged me to "go in the lab and break things" – a reminder to not be afraid to make mistakes in research and to keep pushing the work forward. I learned so much from the members of the Rundell and Voytik-Harbin labs, particularly from my graduate student mentors, Drs. Nimisha Bajaj and Kevin Buno. Their excitement for science kept me motivated and eager to continue as a scientist, even in the face of countless dark hours spent in the confocal microscope room. Nim modeled taking care of yourself and others in graduate school. Kev-o always had a positive outlook even when experiments went wrong. As an undergraduate researcher, I was treated as a colleague and trusted with an independent research project; that was crucial for developing my confidence as a scientist.

I decided to pursue my PhD at the University of Michigan because of the sense of community and the dedication to training, and I am so grateful for the community that supported me throughout my PhD:

I will forever appreciate the efforts of Drs. Eric Martens, Alice Telensnitky, and Denise Kirschner, who helped my partner and me with our "two-body problem" when moving to Michigan. Thank you for valuing me as both a trainee and a person.

The Program in Biomedical Sciences and Office of Graduate and Postdoctoral Studies provides amazing resources to trainees from career advice to conflict resolution to mental

health resources. There is always someone to turn to when you need help, which is so important in graduate school. Thank you to Dr. Scott Barolo, Michelle Melis, Michelle DiMondo, and everyone else in the OGPS office for your efforts on behalf of trainees.

To my cohort: Edmond Atindaana, Yolanda Rivera-Cuevas, Anna-Lisa Lawrence, Austin Campbell, and Matthew Schnizlein. These folks patiently taught me a lot about basic science and microbiology, and I couldn't have made it through our first year courses and Checkpoint 1 without their support. Our honorary cohort member Zack Mendel fostered our community through ping pong and departmental basketball; I thank him for organizing, even though I did break my pinky finger...

Drs. Joyce Wang, Basel Abuaita, and Jay Vornhagen taught me everything I know about wet-lab microbiology.

Drs. Adam Lauring and Krishna Rao provided the opportunity to learn about infectious diseases in the clinic through shadowing.

I have greatly appreciated the mentorship and intellectual insights from members of my committee: Drs. Mary O'Riordan, Mike Bachman, Krishna Rao and Jon Zelner. My research and my training have benefited from their diverse expertise including basic microbiology, epidemiology, clinical microbiology, and infectious disease in the clinic.

I couldn't have asked for a more supportive lab – thank you to my fellow Snitkineers. Everyone in the lab was generous with their time, full of puns, excited about science, had a collaborative spirit, was open to sharing thoughts and code, and always wanted to help improve each others' research and research presentations. I am so thankful that Evan Snitkin took me on as a student and fostered my love for genomics. Our lab "binfie" Ali Pirani had a role in every one of my projects (and truly every project in the lab), and I hope he knows how much he is appreciated by the whole lab. My "work-wife" Zena provided an equal ratio of support to snacks. She was a source of positive energy and the best ideas and was endlessly generous with her time. My undergraduate mentees Geneveive Chiara and Emily Benedict

are thoughtful scientists and thoughtful humans. I am thankful for their contributions to our projects and take solace that these kind women are working in science.

I have learned so much about MRSA epidemiology and MRSA in the clinic from Drs. Kyle Popovich and Bob Weinstein and other collaborators at Rush University Medical Center. This dissertation would not have been possible without their epidemiological study design, metadata and sample collection, and epidemiological and clinical insight. I have absolutely loved working on Team MRSA!

Further, this work would not have been possible without the study volunteers at the Cook County Jail and Cook County Health. Thank you for your contributions to public health, even in a particularly challenging time of your life.

My research and training was supported by the Molecular Mechanisms in Microbial Pathogenesis T32 NIH T32 AI007528, the Rackham Predoctoral Fellowship, and the Rackham Research Grant.

I would like to thank my therapist, who taught me the tools I needed to stay in graduate school, and to all the mental health professionals who support trainees. Anxiety and depression is unfortunately too common among graduate students [1], and I hope this acknowledgement serves as a step towards normalizing seeking help for mental health.

To my family who provided support and encouragement and vocalized often that I could do anything I set my mind to. My parents Vicki and John Thiede demonstrated hard work and service to the community as a teacher and a fireman. In addition to my parents, I was raised by a team of strong women: my sisters Nikki and Samantha, my grandmas Ann and Sharon, and my aunts Mary Ann and Jenny.

I am thankful always to my fiancé Cody Patterson who never lets me forget that I am capable and is always eager to help debug or optimize my code.

Thank you to the musicians who propelled my writing and data analysis – particularly Lorde for releasing Melodrama while I wrote an F31 and Solar Power while I wrote this

dissertation.

And finally, to our cats Jetson, Butterbean, and Jellybean, who provided joy and companionship during the ups and downs of graduate school.

Contents

Dedication ii

Acknowledgements iii

List of Figures xii

List of Tables xiv

Abstract xv

Chapter 1 Introduction 1

1.1 Motivation 1

1.2 Overview of MRSA 2

1.2.1 Origins of MRSA 2

1.2.2 Biology and clinical importance of MRSA 2

1.2.3 Classification of MRSA 3

1.3 MRSA Epidemiology 4

1.3.1 From the hospital to the community and back 4

1.3.2 MRSA in urban communities and incarcerated individuals 5

1.4 Genomic epidemiology to study transmission 6

1.5 Genomic variation leading to emergence and continued success of USA300 . 7

1.6 Datasets 8

1.6.1 Comprehensive MRSA clinical cultures presenting to Cook County Health from 2011-2014 8

1.6.2 MRSA colonization and infection samples among detainees in the Cook County Jail from 2015-2017 . . . 9
1.7 Dissertation Outline . . . 9
Chapter 2 Community Origins of Healthcare-Associated USA300 MRSA Clinical Cultures Revealed with Genomic Epidemiology 10
2.1 Preamble . . . 10
2.2 Introduction . . . 11
2.3 Results . . . 12
2.3.1 Study population and clinical cultures . . . 12
2.3.2 Epidemiological factors associated with onset type . . . 13
2.3.3 Genomic epidemiology across onset type . . . 15
2.3.4 Evidence of potential transmission among infections . . . 16
2.3.5 Healthcare transmission does not drive USA300 MRSA transmission . 16
2.3.6 Individuals use the healthcare system before and after MRSA . . . 18
2.3.7 Individuals with repeat infections tend to be involved in more transmission . . . 19
2.4 Discussion . . . 21
2.5 Methods . . . 23
2.5.1 Study design . . . 23
2.5.2 Statistical tets . . . 24
2.5.3 Whole genome sequencing . . . 24
2.5.4 Identifying putative transmission links . . . 25
2.5.5 Phylogenetic analysis . . . 25
2.5.6 Defining healthcare exposure in the IDPH data . . . 25
2.6 Tables . . . 27
2.7 Supplemental Figures . . . 29

2.8 Funding 30
2.9 Acknowledgements 30
Chapter 3 Genomic Epidemiology of MRSA During Incarceration at a Large, Inner-City Jail **31**
3.1 Preamble 31
3.2 Introduction 32
3.3 Results 33
3.3.1 Features of the study population 33
3.3.2 MRSA colonization patterns during incarceration 33
3.3.3 Clinical MRSA isolates 35
3.3.4 Epidemiologic predictors of MRSA acquisition during incarceration . 36
3.3.5 Genomic epidemiology of USA300 MRSA intake, clinical, and acquisition isolates 36
3.3.6 Relationship of jail location to USA300 MRSA colonization and clinical isolates 38
3.4 Discussion 41
3.5 Methods 44
3.5.1 Study population 44
3.5.2 Swab collection and processing 45
3.5.3 Archived clinical MRSA isolates 45
3.5.4 Whole genome sequencing 46
3.5.5 Location overlap analysis 46
3.5.6 Risk factors and statistical analysis 47
3.6 Supplementary Data 48
3.7 Acknowledgements 48

Chapter 4 Convergent Evolution of a Resistance-Conferring Plasmid in a Large, Urban Jail and the Broader Community **49**

4.1 Preamble 49

4.2 Introduction 49

4.3 Results 50

4.3.1 Transmission among MRSA infections in the Cook County Jail 50

4.3.2 *ermC*-carrying plasmid associated with transmission in Cook County Jail from 2015-2017 52

4.3.3 *ermC*-carrying plasmid present in the larger community 54

4.3.4 *ermC*-carrying plasmid confers constitutive resistance to clindamycin 57

4.3.5 Epidemiological associations with *ermC*-carrying plasmid 57

4.4 Discussion 61

4.5 Methods 63

4.5.1 Data - USA300 MRSA whole-genome sequences in the Cook County Jail and Cook County Health 63

4.5.2 Assessing association of transmission with pangenome 63

4.5.3 Genetic determinants of clindamycin resistance 63

4.5.4 Identifying and comparing plasmid between datasets 64

4.5.5 Epidemiological associations with *ermC*-carrying plasmid 64

4.6 Supplemental Tables 64

4.7 Supplemental Figures 66

Chapter 5 Discussion **68**

5.1 Major dissertation contributions 68

5.1.1 Contributions to our understanding of USA300 MRSA epidemiology . 69

5.1.2 Jails as amplifiers of MRSA transmission in the community 70

5.1.3 A framework for genomic epidemiology in jails 71

5.1.4 The importance of antimicrobial stewardship 73
5.2 The future of genomic epidemiology 74
5.2.1 Moving away from SNV thresholds? 75
5.2.2 Real-time outbreak detection 76
5.2.3 Precision treatment of infection from bacterial genomes 77
5.3 Conclusion 78
Bibliography 79

List of Figures

2.1 Summary of data 13

2.2 Epidemiological factors enriched in HACO- compared to CO- and HO- infection 14

2.3 No distinct sub-lineages causing community and healthcare associated USA300 infections 17

2.4 Exposure to the healthcare system before and after MRSA diagnosis 19

2.5 Analysis of repeat infections 20

2.6 Non-pediatric USA300 clusters 29

3.1 Whole-genome phylogeny of USA300 MRSA infection and colonization isolates in the jail 34

3.2 Comparison of genetic diversity of intake colonization MRSA isolates versus jail-acquired colonization and jail-onset infection 37

3.3 Individuals with closely related USA300 MRSA strains are more likely to reside in common jail locations and have longer length of stay 39

3.4 Overlap in buildings among individuals with closely related MRSA strains . 40

4.1 Phylogenetic tree of USA300 MRSA in Cook County Jail 51

4.2 Association between pangenome and transmission in the Cook County Jail from 2015-2017 53

4.3 Prevalence of *ermC*-carrying plasmid among datasets 55

4.4 Association between *ermC* and transmission in comprehensive clinical cultures from 2011-2014 presenting to Cook County Health 56
4.5 Increasing prevalence and evidence of acquisition of *ermC*-carrying plasmid among repeat infections 59
4.6 Antibiotic exposure in the past 6 months 60
4.7 Combined phylogeny of clinical and jail USA300 MRSA cultures 66
4.8 Pyseer results assessing association between clindamycin resistance and the pangenome in the clinical data 67

List of Tables

2.1 Summary of data 27

2.2 Epidemiological factors associated with onset type 28

3.1 Epidemiological factors associated with acquisition of MRSA colonization . . 44

4.1 Epidemiological associations with harboring the *ermC*-carrying plasmid among the comprehensive sampling of clinical cultures from 2011-2014 58

4.2 Pangenome genes that pass the Bonferonni correction for jail dataset 64

4.3 Pangenome genes that pass the Bonferonni correction for comprehensive clinical culture dataset 65

Abstract

Antibiotic resistant pathogens pose a global health threat as resistance continues to evolve and spread across the globe. One such pathogen, methicillin-resistant Staphylococcus aureus (MRSA) causes about 300,000 cases and 10,000 deaths annually in the United States. MRSA once exclusively caused infections among hospitalized patients with known risk factors. Around 1999, a new lineage called USA300 emerged to cause infection in otherwise healthy individuals in the community in the United States. Since then, infection control efforts have been effective at steadily reducing MRSA rates in the healthcare setting. However, these reductions are strain-specific and MRSA transmission in the community remains a problem. Efforts to reduce community MRSA transmission require in-depth knowledge of how and where MRSA is circulating in the community. In this dissertation, we explored MRSA adaptation and transmission in the urban community of Chicago and in a large, inner-city jail in Chicago.

Leveraging a comprehensive collection of clinical MRSA cultures from 2011-2014, we first applied genomic epidemiology methods to determine the contributions of the community and healthcare settings in MRSA USA300 spread. We found a lack of healthcare overlap among individuals with genetically similar MRSA strains, even among so-called "healthcare-associated" or "hospital-onset" infections. This finding implicated the community in acquisition of USA300 MRSA and merits further studies to pinpoint areas of the community to target.

Next we honed in on one sector of the community with increased burden of MRSA: a

large, urban jail. We found significant importation of MRSA into the jail, with 19% of individuals entering the jail colonized, as compared to the national nasal carriage prevalence of 1.5%. Moreover, by following individuals longitudinally we found that MRSA was also acquired within the jail, with ~8% of intake negative individuals screening positive by 30 days. Further, genomic analysis of individuals acquiring MRSA infections in jail revealed evidence of transmission mediated by spatiotemporal overlap in the jail and persistent environmental contamination, pointing to potential targets to reduce transmission.

Finally, we uncovered an antibiotic resistance-conferring plasmid that was independently acquired multiple times and was associated with increased transmission both in the jail and in the community. The prevalence of the plasmid was much higher in jail-onset infections compared to imported MRSA, suggesting an increased selective pressure in the jail. In the community, the plasmid was associated with incarceration and drug use, suggesting possible acquisition in the jail and potentially reflective of an increased selective pressure for the plasmid in the jail.

In this dissertation, we produced insights into USA300 MRSA epidemiology by applying genomic epidemiology approaches to identify sites and risk factors for community transmission. In particular, our integration of genomic and epidemiological data improved our understanding of transmission pathways in the community, in the jail, and the relationship between the two by identifying a variant selected for in the jail that may be spread to the larger community. Moreover, our delineation of community transmission networks feeding into the jail highlights the need to study additional sectors of the community for MRSA transmission. Finally, we provide an analytical framework for genomic epidemiology in jails, which has subsequently been useful to study transmission dynamics of COVID-19 in the jail population.

Chapter 1

Introduction

1.1 Motivation

Treatment options for bacterial infections have dwindled due to the emergence of antibiotic resistance, with the greatest concern being bacteria that have become resistant to multiple classes of antibiotics and spread across the globe. One such pathogen, methicillin-resistant *Staphylococcus aureus* (MRSA), is a common cause of invasive infection worldwide and deemed a serious threat by the CDC. MRSA was once confined to infecting those with healthcare exposures (hospital-onset or HO-MRSA), but around 1999 a new strain of MRSA called USA300 emerged to cause infections in otherwise healthy individuals in the community (community-onset or CO-MRSA) such as children, incarcerated individuals, and athletes. Subsequently, this strain of MRSA infiltrated the healthcare setting. Successful interventions in the healthcare setting have led to a reduction in HO-MRSA. However, CO-MRSA has not declined to the same extent, in part due to a lack of understanding of the routes of transmission in the community and where interventions would be most impactful in reducing CO-MRSA.

As the most common strain of MRSA is USA300, high-resolution typing via whole-genome sequencing is required to understand routes of spread and to monitor variants circulating and taking hold in the community. We sought to understand the transmission and adaptation of

CO-MRSA in Chicago, IL in one of the largest single-site jails in the United States and in the broader community by integrating genomic and epidemiological data.

1.2 Overview of MRSA

1.2.1 Origins of MRSA

Staphylococcus aureus is both a component of the normal human flora and a successful pathogen [2]. Before antibiotics were available, bloodstream infection with *S. aureus* resulted in 80% mortality [2]. While antibiotics improved patient outcomes, resistance was observed shortly after the introduction of antibiotics to clinical practice – first to penicillin and later to methicillin [3]. Methicillin resistance is thought to have predated methicillin use in the clinic as a result of widespread penicillin use and the subsequent use of methicillin clinically provided a selective pressure for these variants [4]. Resistance is conferred through the acquisition of the mobile genetic element (MGE) SCCmec which occurs in a variety of forms among different strains of MRSA [5].

1.2.2 Biology and clinical importance of MRSA

MRSA can asymptomatically colonize the host for long periods of time; one study observed that ~20% of patients remained colonized 4 years after their documented MRSA colonization or infection [6]. MRSA colonization occurs in about 1.5% of the general population [7] with certain populations such as incarcerated individuals, illicit drug users, homeless individuals, individuals engaging in high-risk sexual behaviors, and those with diabetes, HIV, or on dialysis having higher rates of colonization [8, 9, 10, 11]. In addition to asymptomatic colonization, MRSA can cause a number of infections including skin and soft tissue infections, respiratory, urinary, and bloodstream infections [3]. MRSA colonization is thought to precede

infection and patients are usually infected with their colonizing strain [12].

MRSA causes over 300,000 infections each year and results in over 10,000 deaths in the United States [13] and is also one of the most prevalent pathogens globally [14]. Across multiple studies, MRSA bacteremia has higher mortality rates than methicillin-susceptible *Staphylococcus aureus* MSSA bacteremia [15, 16]. However, there has been inconclusive evidence that MRSA is more virulent than MSSA [16]. Rather, increased mortality may be a result of delay in the microbiologically appropriate treatment or decreased efficacy of vancomycin for treating MRSA infection compared to drugs to treat MSSA [15]. In wound infections, however, similar outcomes exist in MRSA and MSSA [17].

Although there is no conclusive evidence that MRSA is generally more virulent than MSSA, certain MRSA lineages have genetically encoded factors that may increase pathogenic potential [18]. The CO-MRSA strain USA300 and its progenitor, USA500, are particularly pathogenic compared to other CC8 MRSA strains in animal models of bacteremia and abscess [19]. The increase in pathogenicity has been attributed to differential virulence gene expression of core, chromosomally encoded virulence genes regulated by the quorum-sensing gene *agr* [19].

Regardless of questionable differences in virulence, MRSA is challenging to treat as there are less treatment options than MSSA [20]. Moreover, there have been a few cases of development of vancomycin resistance in MRSA, leaving even fewer treatment options [21].

1.2.3 Classification of MRSA

MRSA has traditionally been categorized into molecular types using lower resolution methods than whole-genome sequencing including *spa* typing [22], pulsed-field gel electrophoresis (PFGE) [23] and multilocus sequence typing (MLST) [24]. In this dissertation, I refer to strains as defined by PFGE and MLST. Typing by PFGE is based on the similarity of banding patterns after restriction enzyme digestion. A national database of strains was created with

the naming convention USA followed by numbers, including common strains in the United States such as USA300, USA500, and USA100 [23]. MLST is based on the genetic similarity of seven housekeeping genes [24]; sequence type (ST) 8 encompasses USA300 and USA500 while ST5 encompasses USA100 [3]. Clonal complexes are classified as similarities among 5 of the 7 housekeeping genes with USA300 categorized as CC8. PFGE is labor intensive and classification of banding patterns can be subjective; as such, sequencing probes have been developed to classify strains within CC8 with a PCR assay or in silico sequencing probes [25]. Within these molecular types, there is large genetic variation on the nucleotide level and thus whole-genome sequencing is required to determine transmission relationships between them [25].

1.3 MRSA Epidemiology

1.3.1 From the hospital to the community and back

MRSA was once confined to the hospital setting primarily by the strain USA100, but also by USA200, USA500, and USA800 [23]. Around 1999, one of the the first outbreaks of MRSA in a healthy population with no healthcare exposures occurred in a prison [26] with subsequent reports of MRSA in other correctional facilities [27]. Furthermore, cases were observed in the pediatric population in healthy children [28] and among the St. Louis Rams football team [29]. It was later determined that these community cases were caused by a new strain of MRSA called USA300 by PFGE. Outbreaks continue to happen among children and in daycare centers, in athletes and athletic facilities, jails and prisons, and military barracks as these demographics and places are characterized by one or more of the following: poor hygiene, potential for abrasions, and close person-to-person contact [30, 31]. USA300 is the dominant strain in the community, causing the majority of all skin and soft tissue infections presenting to emergency rooms across the United States [32]. But it has also infiltrated the

healthcare system, having become a common cause of bloodstream infections [33].

In an attempt to predict community or hospital origins of MRSA, MRSA is categorized by time of onset relative to hospitalization with community-onset (CO-MRSA) defined as MRSA whose onset is within 72 hours of hospitalization and hospital-onset MRSA (HO-MRSA) after that threshold. With USA300 now prevalent in both the community and hospital setting, genotype alone (USA100 vs. USA300) does not distinguish site of acquisition. Further, there is evidence that often "hospital-onset" MRSA is due to pre-existing colonization from the community [34]. Thus, higher resolution methods such as whole-genome sequencing are needed to understand the origins of MRSA. We address this in Chapter 2.

1.3.2 MRSA in urban communities and incarcerated individuals

Certain subpopulations of the community are at higher risk for MRSA colonization and infection. Risk factors for MRSA include illicit drug use, HIV, high-risk sexual behaviors, homelessness/alternative housing and incarceration which are factors particularly prevalent in urban communities [35]. These risk factors are highly associated with each other; for example, incarcerated populations are enriched in illicit drug users, HIV infected populations and homlessness which are all also risk factors for MRSA colonization [36, 37].

Incarcerated populations are particularly vulnerable to MRSA colonization, with higher prevalence of colonization than the average MRSA colonization prevalence in the United States [38]. In fact, locations of Cook County with higher rates of MRSA colonization overlap with regions of the community with higher rates of incarceration history [39]. In addition to high risk populations entering the jail, jails are high risk environments for MRSA transmission. Some of the first outbreaks of USA300 were reported in jails and prisons, and outbreaks continue to be a problem.

We and others hypothesize that jails may act as amplifiers for MRSA transmission in the community because there is an intersection of individuals at high risk for MRSA colonization

in an environment that is high risk for MRSA transmission [40]. Upon potentially acquiring MRSA in the jail, many people leave each day and return to their communities, potentially facilitating spread of new MRSA lineages to that community. Also, there is high recidivism in this population, leading to multiple opportunities for acquisition or spread of previously acquired MRSA strains in the jail.

1.4 Genomic epidemiology to study transmission

Genomic epidemiology is an emerging field that combines microbial whole-genome sequencing and epidemiological data to infer transmission networks. As outbreaks of antibiotic resistant pathogens are often caused by epidemic lineages, such as USA300 MRSA, higher resolution methods like whole-genome sequencing are required to discern transmission relationships.

Some of the first genomic epidemiological studies occurred in the healthcare setting to study outbreaks [41] and have subsequently been used to study inter- and intra-healthcare facility transmission [42]. While MRSA outbreaks have been studied in the healthcare setting [43], fewer studies have been conducted in the community. Using genomic epidemiology in the community poses additional challenges as sample and data collection may be challenging, particularly in vulnerable populations such as incarcerated individuals.

In MRSA, one community genomic epidemiological study revealed an important role of households in fueling community MRSA evolution and transmission [44]. A study conducted in the same population as this dissertation used genomics to assess community transmission networks entering a large urban jail and revealed a community transmission network of USA500 MRSA among men who have sex with men (MSM), HIV-positive methamphetamine users [38]. MRSA transmission in jails has been studied with mathematical modeling [45] and risk factors for infection onset in jails assessed with epidemiological studies [46, 47], but to date genomics has not been used to infer MRSA transmission networks within jails. We

use genomic epidemiology in Chapter 3 to assess the extent and routes of MRSA transmission in the Cook County Jail.

1.5 Genomic variation leading to emergence and continued success of USA300

USA300 emerged in the community around 1999 and subsequently has had epidemic success in the United States for decades. The factors that led to USA300's emergence and rapid spread are not fully understood, but acquisition of particular genetic variants including MGEs have been identified as contributors. For example, it is thought that USA300 has enhanced survival on human skin, thus promoting transmission [48]. This may be due to acquisition of the MGE ACME, which is thought to have been acquired by USA300 from S. epidermidis circa 1981 to 1997, around the beginning of the USA300 epidemic [48]. The ACME element contains the gene *speG*, which produces a spermidine acetyltransferase to neutralize polyamines, a byproduct of arginine metabolism in human tissues [48]. Another factor that may have contributed to the evolutionary success of USA300 is the acquisition of fluoroquinolone resistance through chromosomal mutations [49, 50, 51].

In addition to factors that contribute to its broad epidemic success, USA300 may evolve differently in response to varying selective pressures in diverse environments. USA300 is both an important community and hospital pathogen, and likely undergoes different selective pressures in each setting. In fact, recent proteomic studies have shown that USA300 lineages classified as community-associated versus hospital-associated can be distinguished by their proteome [52, 53].

Even within different pockets of the community setting, USA300 MRSA are under differing selective pressures which can lead to different emerging lineages. One example of this is among MSM; USA300 MRSA carrying a plasmid pUSA03 carrying the genes ermC

and mupA conferring resistance to clindamycin and mupirocin respectively were found to be circulating among MSM in Boston and San Francisco [54]. Another example is the varying selective pressures in jails and prisons, which are high transmission settings for infectious disease including MRSA and tuberculosis. Recently, it was shown that highly transmissible, fitness-compensated multidrug-resistant *Mycobacterium tuberculosis* was being selected for in prisons in the country of Georgia resulting in spillover to the community [55]. We wondered if the jail was imposing differing selective pressures in MRSA that facilitate spread, and we address this in Chapter 4.

1.6 Datasets

This dissertation provides insight to MRSA transmission and adaptation associated with increased transmission using two datasets collected in the Cook County Health system including the Cook County Jail.

1.6.1 Comprehensive MRSA clinical cultures presenting to Cook County Health from 2011-2014

The first dataset includes all clinical cultures presenting to Cook County Health from 2011-2014, totalling 1165 cultures from 1101 individuals. It consists mostly of USA300 wound infections. We performed whole-genome sequencing and antibiotic resistance testing on all cultures. Metadata includes demographic information, community exposures including housing status, incarceration and illicit drug use history, binary healthcare exposures including hospitalization, surgery, dialysis, and outpatient and inpatient visits, and exposure to antibiotics in the past 6 months. In addition, we have detailed inpatient and outpatient discharge data from the Illinois Department of Public Health with monthly resolution exposures to 90 inpatient and 96 outpatient facilities across the state of Illinois.

1.6.2 MRSA colonization and infection samples among detainees in the Cook County Jail from 2015-2017

The second data set was collected as part of an epidemiological study in one of the largest single-site jails in the United States: the Cook County Jail in Chicago, IL. Surveillance cultures of MRSA in the nose, throat, and groin were collected among males at intake and after 30 days in jail if the individual remained from January 2016 to December 2017. These colonization isolates, combined with all infections in the jail during the study period and preceding year, underwent whole-genome sequencing. Metadata includes results of a survey of risk factors collected at intake and at day 30 and detailed location data regarding where the individual stayed each day in the jail.

1.7 Dissertation Outline

In the second chapter, I identify origins of MRSA in healthcare-associated community-onset MRSA infections presenting to a large, urban healthcare system in Chicago, IL. In the third chapter, I identify the extent and routes of transmission in a specific section of the community: at entrance to and within a large, urban jail. In the fourth chapter, I identify a plasmid associated with increased transmission in both the jail and larger Chicago community. This dissertation was a result of team science, and I specify my contributions in the preamble of Chapters 2-4.

Chapter 2

Community Origins of Healthcare-Associated USA300 MRSA Clinical Cultures Revealed with Genomic Epidemiology

2.1 Preamble

In this chapter, we assess the contributions of the community and healthcare settings in USA300 MRSA transmission in a comprehensive collection of clinical cultures from Cook County Health across the period of 2011-2014. We use genomic epidemiology to determine the potential origins of MRSA acquisition in the context of the traditional, epidemiological definitions "community-onset", "hospital-onset" and "healthcare-associated community-onset" and find these definitions are not a reliable predictor of acquisition. Across all onset types, there are numerous healthcare exposures, but we find little evidence of healthcare transmission even with detailed inpatient and outpatient discharge data from the Illinois Department of Public Health, suggesting an important role of the community in USA300 MRSA acquisition.

I performed the genomic, phylogenetic, and statistical analyses and created the figures presented in this chapter and drafted this chapter.

2.2 Introduction

Methicillin-resistant Staphylococcus aureus (MRSA) was once confined to causing infections to hospitalized patients with known risk factors [56]. In the United States, healthcare-associated infections were caused predominantly by USA100 MRSA as defined by pulsed-field gel electrophoresis (PFGE) but also by USA200, USA500, USA600 and USA800 [23]. However, around 1999, a new lineage of MRSA called USA300 emerged to infect otherwise healthy individuals with no healthcare exposures in the community including jails and prisons [27, 26], sports teams [29, 57], military barracks [58], and among the pediatric population [28, 59]. Since its emergence, USA300 became the most common cause of skin and soft tissue infections presenting to the emergency room [32]but has also infiltrated the hospital as a common cause of hospital-onset bloodstream infection [33].

In an attempt to predict where MRSA was acquired, MRSA is defined as community-onset (CO) if an infection occurs within 72 hours of hospitalization and hospital-onset (HO) if an infection occurred after that threshold. However, as USA300 is now a common cause of community and healthcare-associated infections, there is a "graying" of what defines hospital or community MRSA [60]. Adding nuance to these epidemiological definitions to capture the role of healthcare exposures in CO-MRSA, the CDC Emerging Infections Program coined the term healthcare-associated community-onset (HACO) wherein an individual had a community-onset infection, but had healthcare exposures in the prior year or previous MRSA [61].

Still, there can be inaccuracies in predicting the site of acquisition by time of infection onset because MRSA can asymptomatically colonize the host for long periods of time [62] and colonization is thought to precede infection [12]. Indeed, genomic epidemiology has revealed that these definitions based on timing of infection onset might inaccurately categorize where MRSA was acquired [34]. Further, exposure to the healthcare setting does not necessarily

mean that MRSA was acquired in that setting [63].

The blurred lines defining these categories manifests when observing the strain-specific trends in MRSA infection prevalence by onset-type from 2005 to 2013; while USA100 MRSA significantly decreased for all three onset types, among USA300, only HO-MRSA cases declined [64, 65]. It is hypothesized that decline of MRSA cases was due to increased awareness and implementation of infection prevention measures to combat catheter-related infections and antibiotic resistant pathogens [64]. These strain-specific differences in response to healthcare interventions could suggest that across all onset categories, USA300 is largely community-acquired, and thus interventions in the healthcare setting have less of an effect on USA300 MRSA cases [65]. Clearly, further work is needed at a higher-resolution than molecular-typing methods (e.g. PFGE) to understand the origins of USA300 MRSA acquisition.

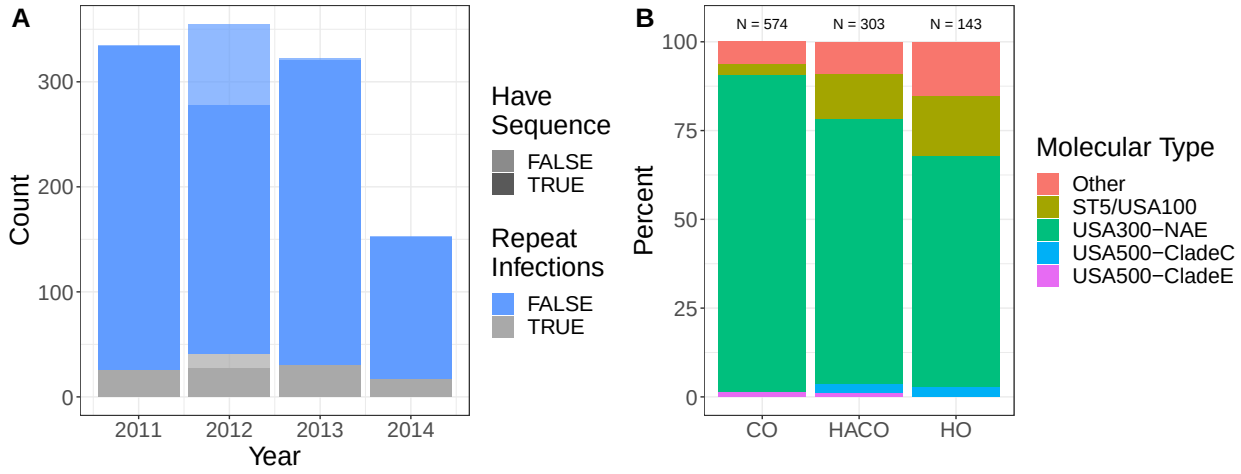
We sought to understand the interplay of community, healthcare, and hospital exposures in USA300 MRSA spread and how these exposures align with onset type definitions. We analyzed a comprehensive collection of MRSA clinical cultures among patients presenting to Cook County Health from 2011-2014, consisting primarily of wound infections, using genomic epidemiology and leveraged a detailed database of hospital discharge data throughout the state of Illinois to directly test the role of the healthcare exposures in MRSA acquisition.

2.3 Results

2.3.1 Study population and clinical cultures

Archived clinical isolates comprising a comprehensive sampling of MRSA infections presenting to Cook County Health over the period were collected from 2011 to 2014 (N = 1203 total samples). The number of clinical MRSA isolates was stable over 2011 to 2013, but declined in 2014 in part because of heuristic diagnosis of skin infections as MRSA and potentially in

Figure 2.1: Summary of data



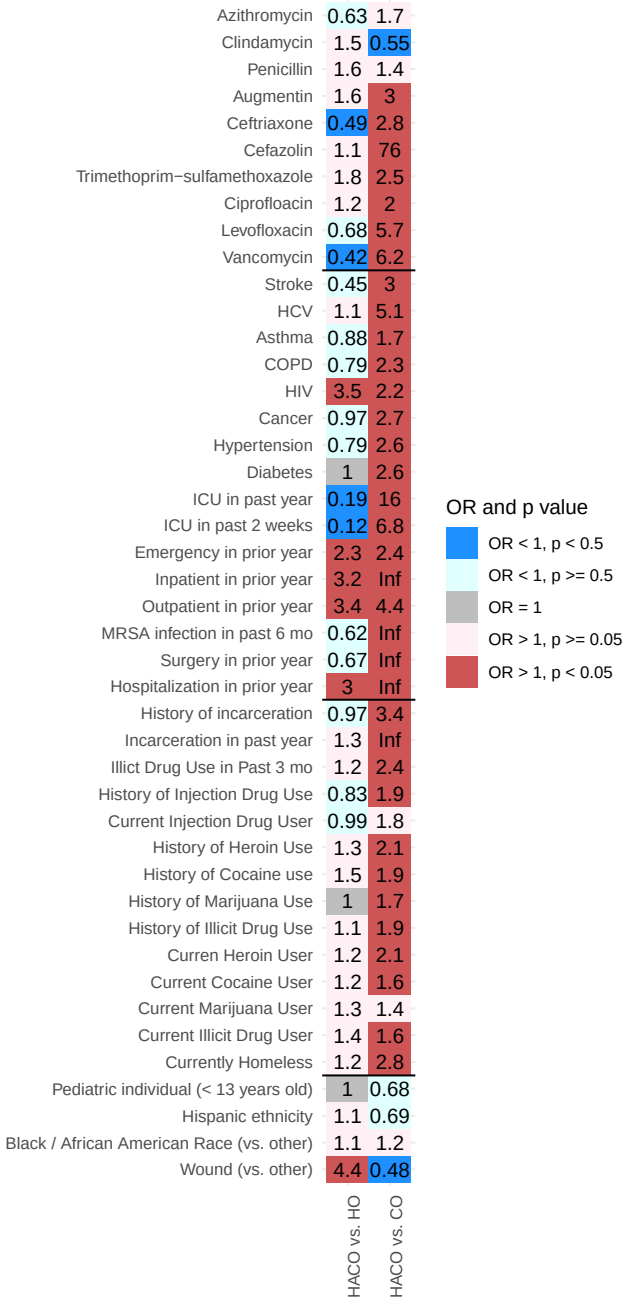
A) Number of clinical cultures over time. Repeat infections are colored in gray. Missing sequences are more transparent blue or gray. B) Per onset type, percent of clinical cultures of each molecular type. USA300-NAE stands for USA300 North American Epidemic lineage, as defined by Bowers et al. [25]. Labeled with total N.

part to declining MRSA cases (Figure 2.1A). There were 113 cases of repeat infections from 49 individuals, defined by $>$ 30 days between clinical samples; 38 samples from 37 individuals were repeated cultures collected less than 30 days apart. Thus, there were a total of 1165 clinical cultures from 1101 individuals over the 4 year time frame. Limiting to samples for which genomes pass quality control and restricting to the first isolate from each patient resulted in a final data set including genomes from 1020 clinical cultures. Overall, wound infections were the most common (81.2%) followed by blood (7.7%) and respiratory (6.2%). Demographic information from these 1020 patients is provided in Table 2.1.

2.3.2 Epidemiological factors associated with onset type

We assessed epidemiological factors among those with HO-MRSA and CO-MRSA compared to those with HACO-MRSA (Table 2.2, Figure 2.2). Compared to HO infections, individuals with HACO infections are enriched in wound cultures, suggesting that HO infections are more

Figure 2.2: Epidemiological factors enriched in HACO- compared to CO- and HO- infection



Visualization of Fisher's exact tests conducted of epidemiological factors enriched in HACO vs. HO and HACO vs. CO. Darker shades indicate significance. Blue indicates OR < 1 and red indicates OR > 1. Grey indicates OR $= 1$. Darker shades indicate significance at $p < 0.05$. Epidemiological factors are broken up by patient characteristics, community factors, healthcare factors including exposures and comorbidities, and antibiotic exposure in the past 6 months. See Table 2 for counts and percentages.

invasive. Individuals with HACO and HO infections have exposures to the healthcare system, but individuals with HO infections have more ICU encounters compared to those with HACO infections, consistent with more severe infection among HO-infected individuals. Individuals with HACO infections have more inpatient, outpatient, and emergency room visits compared to those with HO infections. Individuals with HACO and HO infections have similar levels of community exposures.

By definition, individuals with HACO infections are enriched in healthcare exposures compared to CO infections. As such, antibiotic use is also enriched in individuals with HACO infections compared to those with CO infections. Surprisingly, individuals with HACO infections are enriched with community exposures such as drug use, history of incarceration, and homelesness compared to those with CO infections. In fact, no individuals with CO infections have been incarcerated in the past year of MRSA infections and only 11% have been incarcerated ever compared to 31% and 30% in individuals with HO and HACO infected individuals respectively (Table 2.2).

2.3.3 Genomic epidemiology across onset type

Most cultures were USA300 (N = 832) followed by isolates closely related to USA100/ST5 (N = 80), USA500 including Clades C and E / early branching USA300 (N = 22) and other (N = 86). Onset type classifications were 553 CO, 324 HACO, and 143 HO isolates. USA300 was the most common clinical culture across all onset types. However, a higher percentage of isolates per onset type were USA100 in HO and HACO than in CO, consistent with USA100 as a healthcare-associated pathogen (Figure 2.1B).

Within USA300, the dominant molecular type, we tested if there was intermixing among CO, HO, and HACO isolates on the USA300 phylogeny, or if different sub-lineages of USA300 preferentially spread in community or hospital settings . Previous work by our group focusing on bloodstream infections from 2009-2013 indicated that there were no separate sub-lineages

for CO or HO infections [34]. Expanding our analysis here to a more diverse and comprehensive sampling of clinical isolates supported this previous finding, with onset types showing no evidence of clustering on the USA300 whole-genome phylogeny (Figure 2.3B). Moreover, even when focusing on clusters of isolates that were most closely related, and thereby most closely linked in transmission networks, revealed that most were of mixed onset type. This suggests that transmission networks of USA300 MRSA in the healthcare and the community setting are overlapping.

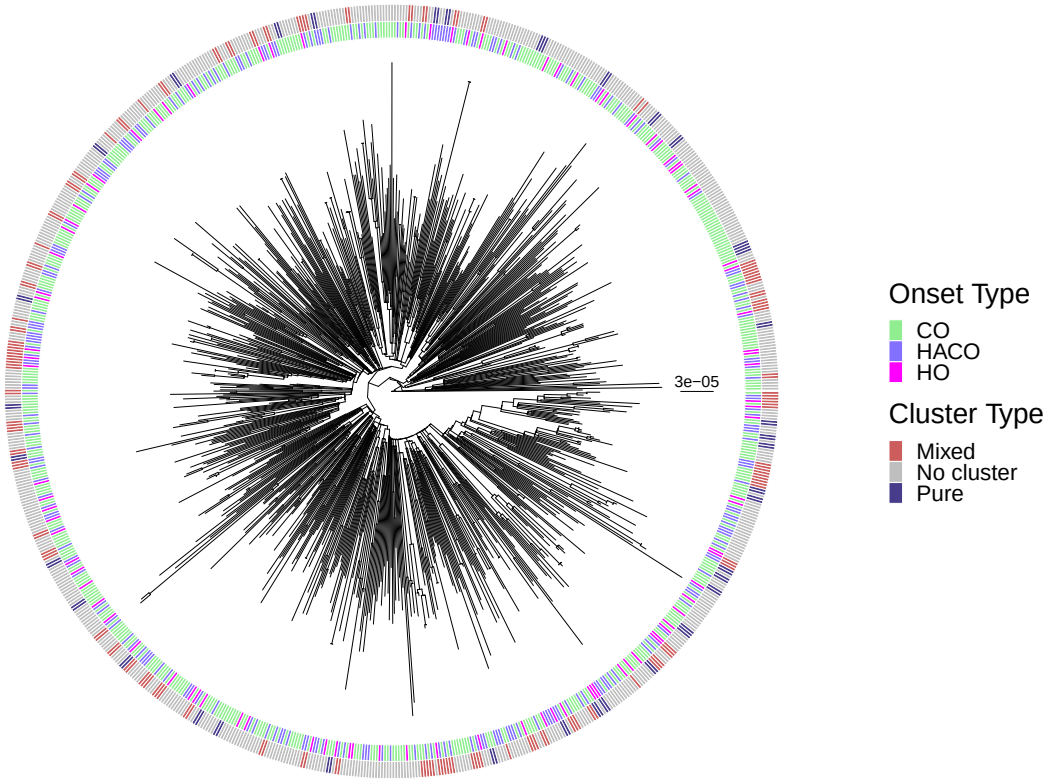
2.3.4 Evidence of potential transmission among infections

We next identified pairs of isolates plausibly involved in direct or indirect transmission by a SNV threshold of 20. We removed individuals less than 13 years of age because of differing epidemiology among adults and children. There were 294 pairs of USA300 isolates making up clusters of size 2-21 (Figure 2.6). It is of note that despite the lack of sampling of asymptomatically colonized individuals, we still identified a large number of individuals potentially related by recent direct or indirect transmission using only clinical cultures. The proportion of CO, HO, and HACO isolates genetically linked to another isolate is similar to the proportion not genetically linked, with HACOs being slightly enriched in transmission clusters ($p = 0.039$) (Figure 2.6.

2.3.5 Healthcare transmission does not drive USA300 MRSA transmission

We next sought to understand what common exposures could be mediating the observed genomic links. We focused on common healthcare exposures as we had access to high resolution healthcare exposure data from the Illinois Department of Public Health with 90 inpatient and 96 outpatient facilities across the state of Illinois from 2013 to 2017. The number of

Figure 2.3: No distinct sub-lineages causing community and healthcare associated USA300 infections



Maximum likelihood phylogeny of USA300 isolates generated by IQ-TREE. One sample per individual. Scale bar individuates substitutions per site. Inner ring indicates onset type. Outer ring indicates if the isolate is in a cluster or not, and if the cluster is of mixed or pure onset type.

genetically linked USA300 pairs where both individuals were diagnosed with MRSA in 2013 or 2014 was 68 out of 294 total pairs. We subsetted our analysis to these genetically linked pairs to reflect the data available from IDPH.

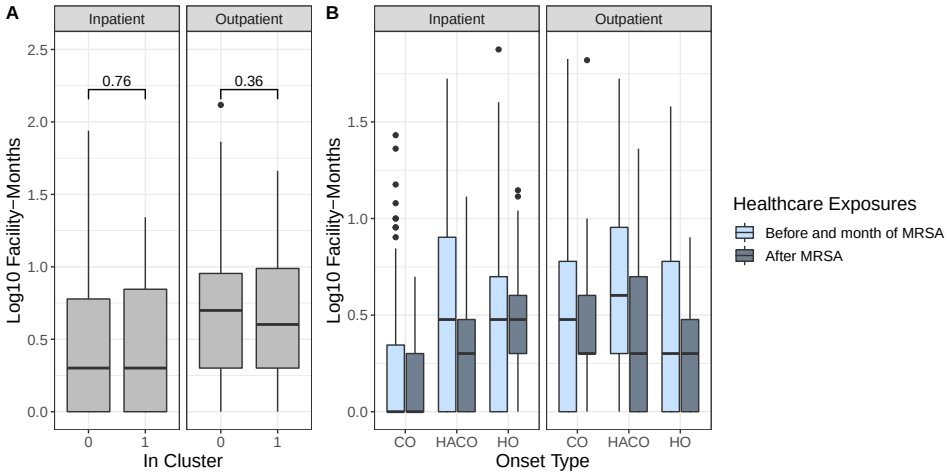
Individuals related to another isolate by 20 SNVs did not have more inpatient or outpatient exposures (Figure 2.4A). We next sought to assess if there was indirect or direct overlap in the healthcare setting among genetically-linked pairs of individuals. Despite high resolution healthcare exposure data, only 5 of 68 pairs overlapped in an inpatient facility in the same month and only 8 of 68 pairs overlapped in an outpatient facility in the same month. All but one of these overlaps occurred in the facility of MRSA diagnosis, which has the most entries, and thus could be random overlap. Indirect overlap was defined by attending the same facility before the date of the latest MRSA diagnosis in the pair. 17 of 68 pairs and 12 of 68 pairs indirectly overlapped in outpatient or inpatient facilities, respectively. Again, all but 2 indirect overlaps occurred at the facility of MRSA diagnosis, and thus could indicate overlap by random chance. Taken with the enrichment of community exposures among HACO and HO infections, this suggests that both HACO and HO infections could be a result of acquisition of colonization outside the healthcare setting.

2.3.6 Individuals use the healthcare system before and after MRSA

Though we found no evidence of extensive transmission in the healthcare setting, individuals do have numerous exposures to the healthcare setting before and after MRSA diagnosis across all onset types (Figure 2.4B) potentially providing opportunities for access to these patients before and after infection. While the median healthcare exposures among CO-MRSA before diagnosis was 0, there were 54 of 196 individuals with CO infections who had an inpatient exposure before or during the month of MRSA diagnosis, indicating that they should be classified as HACO-MRSA. This high resolution data of the broader healthcare network is not typically available to clinicians when making onset type designation. Of note, individuals

within high-risk social networks including those with a history of drug use, incarceration, and homelessness often have exposures to the healthcare setting, including the ER, prior to development of infection (Table 2.2).

Figure 2.4: Exposure to the healthcare system before and after MRSA diagnosis

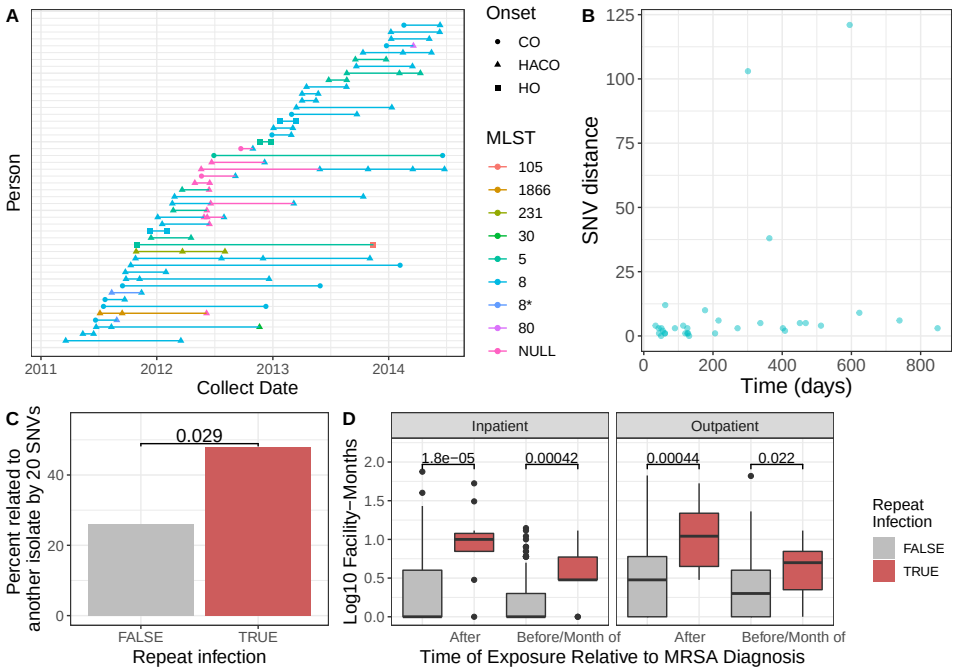


A) Individuals in transmission clusters (In Cluster = 1) do not have more total inpatient or outpatient healthcare exposures indicated by Wilcoxen rank sum test. B) Inpatient and outpatient exposures before and after MRSA by onset type

2.3.7 Individuals with repeat infections tend to be involved in more transmission

There were many repeat infections among individuals and we sought to determine if this was indicative of persistent colonization or acquisition of new strains. There are 108 repeat infections among 47 individuals. Individuals had 2-5 repeat infections during the study period. Of the 47 individuals, 14 had documentation of MRSA in the past 6 months (29.8%), indicating the MRSA culture collected in the study period was not their first infection. While CO-MRSA is the most common infection type in our dataset, HACO-infected individuals are enriched in repeat infections by definition of prior MRSA infection or colonization in the past 6 months (Figure 2.5A).

Figure 2.5: Analysis of repeat infections



A) Individuals with repeat infections, colored by MLST. Shape indicates onset type. B) SNV distance between USA300 repeat infections over time C) Percent of individuals with repeat USA300 infections (N = 23) that are related to another isolate by 20 SNVs compared to those without repeat infection (N = 752). P value from chi-squared test indicates significance. D) Number of healthcare exposures before and month of MRSA diagnosis and after MRSA diagnosis, colored by repeat infection (USA300, non-pediatric,) compared to those with no repeat infection. These are individuals that were infected in 2013 or 2014 (N = 11 repeat infections, N = 331 with no repeat infection). P values from Wilcoxen rank sum tests indicate significantly higher number of exposures among those with repeat infections across all inpatient, outpatient, before and after MRSA.

40 of the 47 individuals have multiple genomes available across repeat infections. Of those 40 individuals, 34 (85%) have an infection of the same MLST over time, with 23 having repeat USA300 infections over time. The majority of individuals with repeat USA300 infections have an infection of the same strain; only 3 of 24 individuals had at least 1 USA300 repeat infection that was greater or equal to 20 SNVs from a previous infection (Figure 2.5B). Individuals had a repeat infection with the same strain of MRSA (within 20 SNVs) up to 849 days apart (Figure 2.5B). This indicates long-term persistent colonization or constant

exposure to an environmental source. Individuals with repeat infections have 2.6 times higher odds of being related to another isolate by 20 SNVs than those without repeat infections ($p = 0.029$) (Figure 2.5C). This suggests a crucial role for individuals with persistent colonization in transmission networks. Furthermore, individuals with repeat infections tend to have more healthcare exposures before and after MRSA (Figure 2.5D) suggesting these individuals are frequent users of the healthcare system. It is possible that they acquired their strain in the healthcare system, although we have little evidence of this occurring. This higher exposure to the healthcare system could also suggest underlying comorbidities that result in persistent colonization and repeat infection.

2.4 Discussion

Infection prevention efforts and antimicrobial stewardship in the healthcare setting have made significant strides in reducing MRSA infection rates [64], but this progress has mainly occurred in USA100 MRSA and less so in USA300 MRSA [65]. We aimed to understand where USA300 MRSA acquisition is occurring and how it relates to commonly-used onset type definitions. From the lack of healthcare overlap and evidence of persistent colonization manifesting in repeat infection with the same strain, it is clear that location of onset or recent healthcare exposure are not sufficient to attribute the hospital as a source of infection among USA300 MRSA. This calls into question the utility of these epidemiological onset type definitions in predicting sites of MRSA acquisition.

Consistent with previous work, we showed that there are no separate lineages of USA300 MRSA circulating in the healthcare and community settings [34]. We observed community exposures among healthcare-associated infections and a lack of healthcare overlap among transmission pairs indicating that acquisition of USA300 in the community often manifests as HACO- and HO- infections. This is consistent with the smaller impact of infection control

to reduce USA300 infection compared to USA100 in the healthcare setting [65].

While we observed no evidence of USA300 healthcare transmission, individuals with MRSA infection have exposures to the inpatient and outpatient setting before and after MRSA. Individuals in high-risk social networks such as those with incarceration history, those who use illicit drugs, and homeless individuals are frequent users of the healthcare system. MRSA surveillance or education on the risk of MRSA and the need for enhanced personal hygiene and environmental cleaning could be implemented during a healthcare visit to attempt to prevent infection.

Further, individuals with repeated infection, suggestive of persistent colonization, are involved in more putative transmission clusters. Targeting these individuals for interventions could be a potential intervention for reducing MRSA transmission. Importantly, these individuals use the healthcare system more before and after diagnosis, indicating potential points of intervention and as such decolonization may be an effective strategy for this population. A multi-center, randomized control trial conducted by Huang et al showed that decolonization with CHG and nasal mupirocin at discharge reduced repeated MRSA infections [66].

Our study was limited in that we only collected clinical cultures and not surveillance cultures. Despite collecting only clinical cultures, we were able to detect many putative transmission links. The stark lack of healthcare transmission could be due to missing intermediates in the transmission network who are asymptomatically colonized. However, in a previously published study conducted at the Cook County Jail, we again observed many genetic linkages among infected individuals but were able to detect significant location overlap explaining transmission [67], suggesting that at least there is less transmission in the healthcare network than in the confined setting of a jail. Further, we only collected healthcare exposure data at monthly and facility resolution. However, even if higher resolution data was available, there would be only a few number of pairs to assess at a more granular level. Varying definitions for HACO exist, some with [61, 68] and some without [64] the inclusion

of prior MRSA colonization and infection in the definition. In our data, 11.6% (N = 32) of USA300 isolates classified as HACOs were classified as such solely because of a prior MRSA infection. Regardless of how HACO is defined, our data suggest that neither a history of MRSA or healthcare exposure equate to healthcare acquisition.

As we have shown that most USA300 infections seem to be acquired outside of the healthcare setting, further work to understand how transmission is occurring in the community is warranted, particularly for those with no known risk factors for MRSA. In the past, we have shown that there are high rates of MRSA at community infectious disease clinics [38] and among HIV infected individuals, high burden of MRSA at entrance to jails [38], and high rates of MRSA acquisition within jails [67]. More studies using epidemiological data and whole-genome sequencing are crucial to understand community reservoirs of MRSA and routes of transmission and to move the needle on USA300 infections.

2.5 Methods

2.5.1 Study design

We examined existing clinical MRSA isolates from 2011-2014 isolated from patients seeking care at Cook County Health, the major public healthcare network in Chicago, IL. Comprehensive sample collection was conducted during these years. We performed electronic and manual chart review to ascertain community (e.g., unstable housing, illicit drug use, incarceration history), demographic information, healthcare exposures, and comorbidities for included individuals. Outpatient and inpatient visits from discharge data in the state of Illinois from 2013-2017 were queried by the Illinois Department of Public Health (IDPH).

Isolates were defined as hospital-onset (HO) if MRSA onset was 72 hours after hospitalization. Isolates were considered healthcare-associated community-onset (HACO) if onset was within 72 hours of hospitalization and the individual had prior healthcare exposure in-

cluding hospitalization, surgery, dialysis, long term care in the past year or MRSA infection or nares colonization in prior 6 months and community-onset (CO) if they did not have these exposures.

Repeat infections were defined as having a repeat culture more than 30 days after the first culture, otherwise these were repeat cultures and were excluded from the analysis. We defined pediatric as less than 13 years of age.

2.5.2 Statistical tets

We compared epidemiological factors enriched in HACO-infected individuals compared to HO- and CO- infected individuals using two independent Fisher's exact tests using the exact2x2 package in R. Null data was removed and denominators are specified. Infection type, ethnicity, and race were binarized. We compared time in facilities with a Wilcoxen rank sum test using the base R function wilcox.test.

2.5.3 Whole genome sequencing

Genomic DNAs extracted from MRSA isolates were prepared for sequencing using a Nextera XT library preparation kit (Illumina, San Diego, CA) or NEBNext Ultra (Illumina, San Diego, CA) Library Preparation kit according to manufacturer instructions. Sequencing was performed on an Illumina NextSeq500 or Illumina NovaSeq instrument using a high-output kit with paired-end 2x75 or 2x150 base reads, respectively. Library preparation and sequencing were performed at the Microbiome Core and Advanced Sequencing Core, respectively at the University of Michigan. Variant calling pipeline can be found on Github: https://github.com/Snitkin-Lab-Umich/variant_calling_pipeline Raw sequence data was deposited under Bioproject PRJNA734638. We conducted multilocus sequence typing using ARIBA [69] and further classified CC8 isolates using in silico sequencing probes provided by

Bowers et al [25]. The size of the USA300 core genome was 2.58Mb.

2.5.4 Identifying putative transmission links

We classified individuals that are plausibly involved in recent transmission based on a SNV threshold of 20 to prioritize minimizing false positives [70, 71]. Additionally, most individuals with repeat USA300 infections had subsequent infections with a strain within 20 SNVs of the first infection.

2.5.5 Phylogenetic analysis

After generating the whole genome alignment, we then masked sites identified as recombinant by Gubbins(Croucher et al., 2015) and used this masked whole-genome alignment to build a maximum likelihood phylogeny with IQ-TREE [72, 73]. Non-USA300 tips were dropped using the drop.tip function in ape [74]. We overlayed metadata on phylogenetic tree using ggtree [75], gheatmap, and ggnewscale.

2.5.6 Defining healthcare exposure in the IDPH data

The IDPH data contained monthly level resolution of exposures to inpatient and outpatient facilities. We defined exposure to a facility in terms of "facility-months" where a stay from Jan 2013 to Jan 2013 would be recorded as 1 facility-month (typical of an outpatient exposure) and a stay from Jan 2013 to March 2013 would be recorded as 3 facility-months (more typical of an inpatient exposure). For visualization purposes, we added 1 and log10 transformed the data, such that a log10 value of 0 indicates no healthcare exposures. We aggregated exposures by before and the month of MRSA and after MRSA; note that with monthly level resolution, an exposure occurring the month of MRSA could have happened before or after the MRSA diagnosis or might represent the visit where MRSA was diagnosed.

The study was approved with waiver of consent by the Cook County Institutional Review Board.

2.6 Tables

Table 2.1: Summary of data

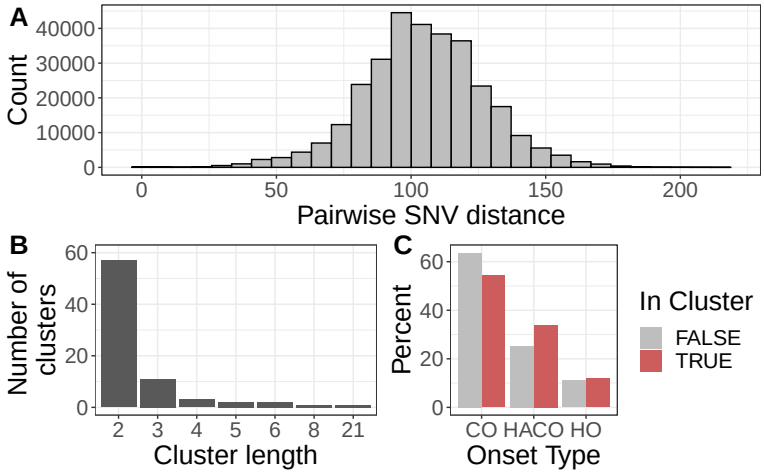
Variable	Number (Percent)
RACE	
African American/Black	612 (60)
White	319 (31.27)
American Indian/Alaska Native	57 (5.59)
Asian	21 (2.06)
Other/UTD (unable to determine)	6 (0.59)
Multiple	3 (0.29)
Native Hawaiian/Pacific Islander	2 (0.2)
ETHNICITY	
Non-Hispanic/Latino/Spanish Origin	792 (77.65)
Hispanic/Latino/Spanish Origin	222 (21.76)
Unknown	2 (0.2)
ONSET TYPE	
CO	574 (56.27)
HACO	303 (29.71)
HO	143 (14.02)
INFECTION TYPE	
Wound	828 (81.18)
Blood	79 (7.75)
Respiratory	63 (6.18)
Fluid	22 (2.16)
Other	16 (1.57)
Urine	12 (1.18)

Table 2.2: Epidemiological factors associated with onset type

	Percent epidemiological factor			HACO vs. HO			HACO vs. CO		
	HO	HACO	CO	p	OR	95% CI	p	OR	95% CI
DEMOGRAPHIC									
Wound (vs. other)	53% (50/94)	84% (193/231)	91% (469/513)	4.1E-08	4.4	2.6,7.6	0.0022	0.48	0.3,0.77
Black / African American Race (vs. other)	61% (57/94)	64% (147/231)	59% (301/513)	0.62	1.1	0.67,1.9	0.22	1.2	0.89,1.7
Hispanic ethnicity	18% (16/91)	19% (43/231)	25% (127/512)	0.87	1.1	0.57,2.1	0.073	0.69	0.47,1
Pediatric individual (<13 years old)	4.3% (4/94)	4.3% (10/231)	6.2% (32/513)	1	1	0.3,3.6	0.39	0.68	0.31,1.4
COMMUNITY FACTORS									
Currently Homeless	11% (10/92)	13% (30/229)	5.1% (26/506)	0.71	1.2	0.56,2.8	0.00044	2.8	1.6,5
Current Illicit Drug User	36% (32/90)	43% (98/227)	32% (150/475)	0.25	1.4	0.82,2.3	0.0031	1.6	1.2,2.3
Current Marijuana User	20% (18/90)	25% (56/227)	19% (89/474)	0.46	1.3	0.72,2.5	0.074	1.4	0.96,2.1
Current Cocaine User	14% (13/90)	17% (39/227)	11% (53/474)	0.62	1.2	0.61,2.5	0.032	1.6	1,2.6
Curren Heroin User	18% (16/90)	21% (48/227)	11% (53/474)	0.54	1.2	0.64,2.4	0.00078	2.1	1.4,3.3
History of Illicit Drug Use	58% (53/91)	60% (137/228)	44% (211/477)	0.8	1.1	0.64,1.8	0.00011	1.9	1.4,2.6
History of Marijuana Use	40% (36/91)	40% (92/228)	29% (138/477)	1	1	0.62,1.7	0.0034	1.7	1.2,2.3
History of Cocaine use	23% (21/91)	31% (71/228)	19% (92/477)	0.17	1.5	0.85,2.7	0.00059	1.9	1.3,2.7
History of Heroin Use	23% (21/91)	28% (63/228)	16% (74/477)	0.48	1.3	0.71,2.3	0.00023	2.1	1.4,3.1
Current Injection Drug User	8.9% (8/90)	8.8% (20/228)	5.1% (24/475)	1	0.99	0.42,2.4	0.067	1.8	0.96,3.4
History of Injection Drug Use	15% (14/91)	13% (30/228)	7.3% (35/477)	0.59	0.83	0.42,1.7	0.017	1.9	1.1,3.2
Illict Drug Use in Past 3 mo	23% (21/90)	26% (60/227)	13% (62/477)	0.67	1.2	0.65,2.2	1.8E-05	2.4	1.6,3.6
Incarceration in past year	19% (18/94)	24% (55/231)	0% (0/513)	0.38	1.3	0.73,2.5	8.1E-31	Inf	44,Inf
History of incarceration	31% (29/94)	30% (70/231)	11% (58/513)	1	0.97	0.57,1.6	1E-09	3.4	2.3,5.1
HEALTHCARE FACTORS									
Hospitalization in prior year	49% (46/94)	74% (171/231)	0% (0/513)	2.7E-05	3	1.8,4.9	3.3E-117	Inf	350,Inf
Surgery in prior year	56% (53/94)	46% (107/231)	0% (0/513)	0.11	0.67	0.41,1.1	2E-64	Inf	110,Inf
MRSA infection in past 6 mo	38% (36/94)	28% (64/231)	0% (0/513)	0.065	0.62	0.36,1	3.2E-36	Inf	48,Inf
Outpatient in prior year	44% (41/94)	73% (168/231)	37% (192/513)	1.1E-06	3.4	2.1,5.7	2.3E-19	4.4	3.1,6.3
Inpatient in prior year	50% (47/94)	76% (176/231)	0% (0/513)	6.6E-06	3.2	1.9,5.3	3.5E-122	Inf	390,Inf
Emergency in prior year	36% (34/94)	57% (132/231)	36% (183/513)	9E-04	2.3	1.4,3.9	6.6E-08	2.4	1.7,3.3
ICU in past 2 weeks	18% (17/94)	2.6% (6/231)	0.39% (2/513)	4.5E-06	0.12	0.046,0.32	0.013	6.8	1.3,47
ICU in past year	26% (24/94)	6.1% (14/231)	0.39% (2/513)	3.8E-06	0.19	0.09,0.4	3.7E-06	16	3.9,100
Diabetes	36% (34/94)	37% (85/231)	18% (93/513)	1	1	0.62,1.7	8.8E-08	2.6	1.9,3.7
Hypertension	57% (54/94)	52% (119/231)	29% (148/513)	0.39	0.79	0.48,1.3	3.8E-09	2.6	1.9,3.6
Cancer	19% (18/94)	19% (43/231)	7.8% (40/513)	1	0.97	0.51,1.9	4.3E-05	2.7	1.7,4.4
HIV	4.3% (4/94)	13% (31/231)	6.6% (34/513)	0.017	3.5	1.2,11	0.0032	2.2	1.3,3.7
COPD	12% (11/94)	9.5% (22/231)	4.3% (22/513)	0.55	0.79	0.36,1.8	0.007	2.3	1.2,4.4
Asthma	23% (22/94)	21% (49/231)	14% (71/513)	0.66	0.88	0.49,1.6	0.013	1.7	1.1,2.5
HCV	9.6% (9/94)	11% (25/231)	2.3% (12/513)	0.84	1.1	0.51,2.6	3.7E-06	5.1	2.5,11
Stroke	7.4% (7/94)	3.5% (8/231)	1.2% (6/513)	0.15	0.45	0.16,1.4	0.042	3	1,8.8
ANTIBIOTIC EXPOSURE IN PAST 6 MO									
Vancomycin	74% (70/94)	55% (127/231)	16% (84/513)	0.0011	0.42	0.24,0.73	4.8E-26	6.2	4.4,8.9
Levofloxacin	20% (19/94)	15% (34/231)	2.9% (15/513)	0.25	0.68	0.36,1.3	1.8E-08	5.7	3,11
Ciprofloacin	8.5% (8/94)	10% (23/231)	5.3% (27/513)	0.84	1.2	0.49,2.9	0.026	2	1.1,3.6
Trimethoprim-sulfamethoxazole	17% (16/94)	27% (62/231)	13% (65/513)	0.064	1.8	0.97,3.3	4.7E-06	2.5	1.7,3.8
Cefazolin	12% (11/94)	13% (30/231)	0.19% (1/513)	0.85	1.1	0.53,2.4	3.3E-15	76	13,1500
Ceftriaxone	27% (25/94)	15% (35/231)	6% (31/513)	0.019	0.49	0.27,0.9	0.00013	2.8	1.7,4.7
Augmentin	8.5% (8/94)	13% (30/231)	4.7% (24/513)	0.34	1.6	0.71,3.9	0.00011	3	1.7,5.3
Penicillin	2.1% (2/94)	3.5% (8/231)	2.5% (13/513)	0.73	1.6	0.35,11	0.48	1.4	0.55,3.4
Clindamycin	37% (35/94)	48% (110/231)	62% (319/513)	0.11	1.5	0.93,2.5	0.00022	0.55	0.4,0.76
Azithromycin	15% (14/94)	10% (23/231)	6% (31/513)	0.25	0.63	0.3,1.3	0.067	1.7	0.97,3

2.7 Supplemental Figures

Figure 2.6: Non-pediatric USA300 clusters



A) Pairwise core SNV distance distribution. Core genome size is 2.5Mb. B) Cluster size distribution based on SNV threshold of 20 SNVS. C) Percent CO, HACO, and HO isolates overall and only those that are related to another isolate by 20 SNVs. Slight differences exist between the number of CO, HO, and HACO in a cluster versus not, with more HACOs and less COs being related to another isolate by 20 SNVs by a three-way chi-square test ($p = .039$).

2.8 Funding

Funding from CDC Broad Agency Announcement: Genomic Epidemiology of Community-Onset Invasive USA300 MRSA Infections; Contract ID: 75D30118C02923.

2.9 Acknowledgements

Contributors to this work include: William Trick, Alla Aroutcheva, Michael Schoeny, Robert Weinstein, Evan Snitkin and Kyle Popovich.

Chapter 3

Genomic Epidemiology of MRSA During Incarceration at a Large, Inner-City Jail

3.1 Preamble

We and others hypothesize that jails may amplify MRSA transmission in the community: there is a high burden of individuals colonized with MRSA entering the jail from diverse communities [38], and the jail provides an opportunity for these individuals to interact. With ~250 individuals entering and exiting the jail everyday and a daily census of 9000-10000, this could provide an opportunity for individuals to introduce strains that they acquired in jail back to their communities upon release. This chapter uses genomic epidemiology to understand the extent and routes of transmission in the Cook County Jail in Chicago, IL which is one of the largest single-site jails in the country. The following chapter explores the biological underpinnings of the transmission we observed in the jail.

This work was published in Clinical Infectious Diseases in January 2021:

Popovich, Kyle J, Stephanie N Thiede, Chad Zawitz, Alla Aroutcheva, Darjai Payne, William Janda, Michael Schoeny, Stefan J Green, Evan S Snitkin, and Robert A Weinstein. "Genomic Epidemiology of MRSA During Incarceration at a Large Inner-City Jail." Clinical Infectious Diseases, no. ciaa1937 (January 4, 2021). https://doi.org/10.1093/cid/ciaa1937.

I am a co-first author on this paper with Dr. Popovich. I performed genomic, phylo-

genetic, and location overlap analyses, created figures and wrote supplemental materials, assisted with interpretation of results, and assisted with drafting the manuscript.

3.2 Introduction

Methicillin-resistant *Staphylococcus aureus* (MRSA) is a significant cause of clinical infection in urban communities [56]. Congregate living (homeless shelters, military barracks, and correctional facilities), close person-to-person contact, sharing personal items, poor hygiene, environmental contamination, and compromised skin integrity promote MRSA transmission [76]. Although infection control recommendations address these factors [77], certain community settings provide special challenges.

Correctional facilities, jails and prisons, are congregate settings where outbreaks of MRSA have occurred [78, 27, 79]. In contrast to prison, jails have relatively short-term incarcerations while detainees await sentencing, with high turnover and recidivism. These features could augment MRSA spread. Models suggest that MRSA transmission occurs during incarceration; individuals colonized with MRSA are the primary source of transmission; and following discharge, in the absence of jail interventions to control MRSA, resistance spreads to the community at large [80]. However, lack of data on MRSA transmission in jails has hampered establishing these facilities as potential key points of intervention.

A complicating factor in understanding transmission dynamics of MRSA and identifying targets for interventions is that individuals entering jail may be colonized due to high-risk community exposures [81, 82, 38], illicit drug use, unstable housing, and type and location of residence [83, 39, 40, 54, 84]. Such individuals may then be at risk for developing MRSA infection during incarceration and have the opportunity to intermingle with other individuals, potentially increasing MRSA spread.

Prior work in urban jails demonstrated that jail-based interventions can significantly

impact community disease patterns (eg, sexually transmitted diseases) [85, 86, 87, 88]. It remains unclear if urban jails are nonhospital settings that are a controllable focus of MRSA and if a jail intervention could have downstream benefit to the community-at-large for reducing MRSA burden. Therefore, the study objectives were to (1) examine the rate of MRSA acquisition during incarceration at a large urban jail, (2) identify epidemiologic and jail-based predictors of MRSA acquisition, and (3) characterize the genomic epidemiology of colonizing and clinical MRSA strains.

3.3 Results

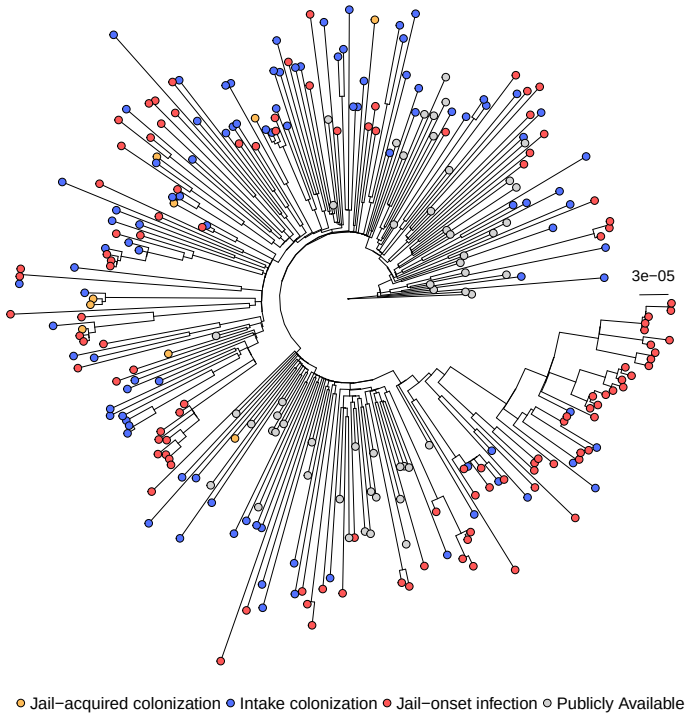
3.3.1 Features of the study population

There were 718 unique individuals (800 incarcerations) enrolled. The prevalence of MRSA colonization at intake was 19% [10]. Strains brought into the jail by those colonized at intake were diverse (Supplementary Figures S5 and S6). Among those enrolled, 267 (33%) incarcerations lasted 30 days or longer. Of those remaining incarcerated at Day 30, 160 individuals accounting for 184 (70%) incarcerations completed the Day 30 study visit. Among those completing the Day 30 study visit, 82% were African-American and 7% Hispanic. Use of illicit drugs before incarceration was common among individuals who completed the Day 30 study visit, with 80% reporting use in the past year. Recidivism was high; 91% of individuals in the study had prior jail incarceration.

3.3.2 MRSA colonization patterns during incarceration

Of the 184 detainees with a completed Day 30 study visit, 41 (22%) were positive and 143 (78%) were negative for MRSA at admission. Of the 143 negative at admission, 131 (91.6%) remained negative at the Day 30 study visit and 12 (8.4%) acquired MRSA (Supplementary

Figure 3.1: Whole-genome phylogeny of USA300 MRSA infection and colonization isolates in the jail



Recombination-masked whole-genome alignment was used to make a maximum likelihood phylogeny of intake colonization, jail-acquired colonization, and jail-onset infection collected from individuals in the jail and publicly available genomes [25]. Tree is midpoint rooted. For samples in the current study, only a single isolate per individual was included, unless genomic analysis supported multiple isolates from an individual being associated with independent acquisition events (see Methods). Overall, jail isolates span the full diversity of the USA300 phylogeny, with intermixing of in- take colonization, jail-acquired colonization, and jail-onset infection isolates. However, in the background of this diversity, clustering of isolates can be observed, particularly for jail-onset infections. Publicly available isolates span the diversity of the tree, but do not interrupt clusters of jail samples. Scale bar represents substitutions per site. One isolate with a long branch was removed for visualization purposes (see full USA300 tree in Supplementary Figure S8).

Figure S7). Of the 12 acquisitions, 2 were the same individual who, in sequential incarcerations, separated by 3 months, was negative at intake but colonized at Day 30. For this individual, the putatively-acquired strains were >1000 SNVs apart, supporting acquisition of a new strain rather than intermittent carriage. Of the 41 incarcerations positive at admission, 17 (41%) were no longer colonized, and 24 (59%) remained colonized at the Day 30 study visit (Supplementary Figure S7). One of these 24 detainees acquired a new strain of MRSA in the nares by the Day 30 visit and maintained throat colonization with the initial strain.

For the 12 MRSA acquisitions, 9 (75%) individuals were colonized at 1 body site, 2 (17%) at 2 body sites, and 1 (8%) at 3 body sites. By body site, 2 (17%) individuals had throat colonization, 7 (58%) nares colonization, and 7 (58%) inguinal colonization. There were no differences in likelihood of acquisition of colonization by body site.

For the 24 participants who had persistent colonization, 8 (33%) were colonized at 1 body site, 6 (25%) at 2 body sites, and 10 (42%) at 3 body sites. By body site, 18 (75%) had throat colonization, 16 (67%) nares colonization, and 16 (67%) inguinal colonization. There were no differences in likelihood of persistent colonization by body site. Individuals colonized at multiple body sites were usually colonized with the same strain (Supplementary Figure S2).

3.3.3 Clinical MRSA isolates

There were 142 representative clinical MRSA isolates from male detainees who underwent whole genome sequencing (WGS); 125 were jail-onset during the study period, 3 were community- onset during the study period, and 14 were isolated in the year prior. Sequenced clinical isolates were mostly from skin/skin structure infections (98.6%) and identified as USA300 (92%). Of individuals with clinical infections, 66% were African- American, 8.5% Hispanic; mean age was 37 (SD 12) years; 43% reported current illicit drug use; and

11.3% were living on the street before incarceration.

3.3.4 Epidemiologic predictors of MRSA acquisition during incarceration

Among the 12 individuals who acquired MRSA colonization, 11 were African-American and none Hispanic. Among exposures before incarceration (Table 1), heroin use was significantly associated with acquiring MRSA colonization (P = .05). No other types of drug use were associated with acquisition. While HIV status was not associated with MRSA acquisition, taking antiretrovirals was negatively associated with MRSA acquisition (P = .08)

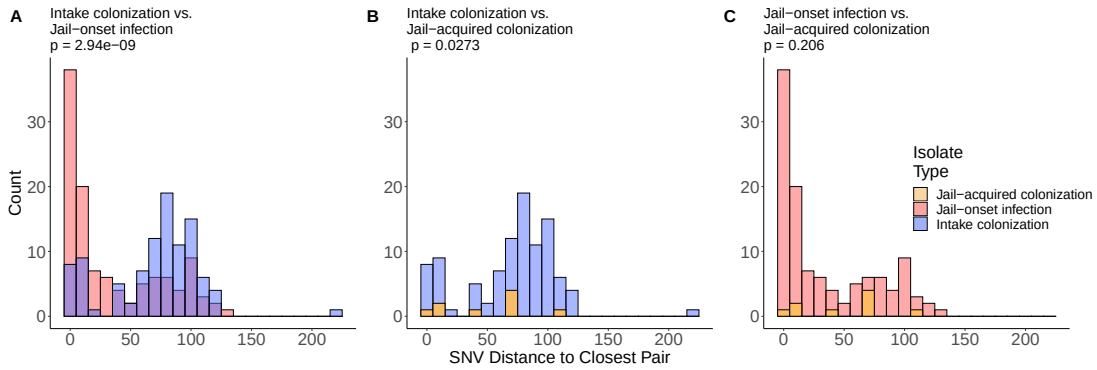
Among exposures occurring during incarceration, sharing personal items was significantly associated with MRSA acquisition (OR 4.92; 95% CI: 1.45, 16.67, P = .01). A variety of personal items were shared and no one individual item was associated with increased risk of MRSA acquisition.

3.3.5 Genomic epidemiology of USA300 MRSA intake, clinical, and acquisition isolates

While USA300 intake, jail-onset clinical, and colonization acquisition isolates were overall diverse, clusters of closely related strains were identified (Figure 3.1, Supplementary Figure S8). Examining strains with close genetic neighbors revealed in comparison to intake isolates, jail-onset USA300 clinical and acquisition colonization isolates were more likely to have closely related genetic neighbors (Figure 3.2), suggesting the existence of transmission networks that included both colonized and infected individuals. Significance remained when adjusting for difference in sample sizes between the potential sources (Supplementary Figure S9).

Four acquisition isolates were closely related to at least one other isolate (range 1–3) (4

Figure 3.2: Comparison of genetic diversity of intake colonization MRSA isolates versus jail-acquired colonization and jail-onset infection



To evaluate whether jail-acquired USA300 MRSA colonization and jail-onset USA300 MRSA infections were enriched for recent transmission events, their genetic diversity was compared to that of intake USA300 MRSA colonization by creating distributions of genetic distances to closest genetic neighbors (core genome size = 2.54 Mb). Closest-pair sources for intake USA300 MRSA colonization include other intake isolates (n = 100). Closest-pair sources for jail-acquired USA300 MRSA colonization (n = 9) and jail-onset USA300 MRSA infections (n = 113) included all isolate types (n = 239 sources including intake positive colonization, jail-onset infection, jail-acquired colonization, community-onset infection (n = 3), infections that occurred in 2015 but were in jail during the study period (n = 14)). Wilcoxon rank-sum test was used to make pairwise comparisons between the 3 sets— one-sided test for A and B, two-sided test for C. Comparisons are shown for (A) jail-onset infections versus intake colonization, (B) jail-acquired colonization versus intake colonization, and (C) jail-acquired colonization versus jail-onset infections. Histograms are overlapping, not stacked, and colors are blended in overlapping parts of distributions. Significance remained when controlling for the differences in number of possible pairs for intake colonization and jail-onset infections (see Supplementary Methods, Supplementary Figures S9).

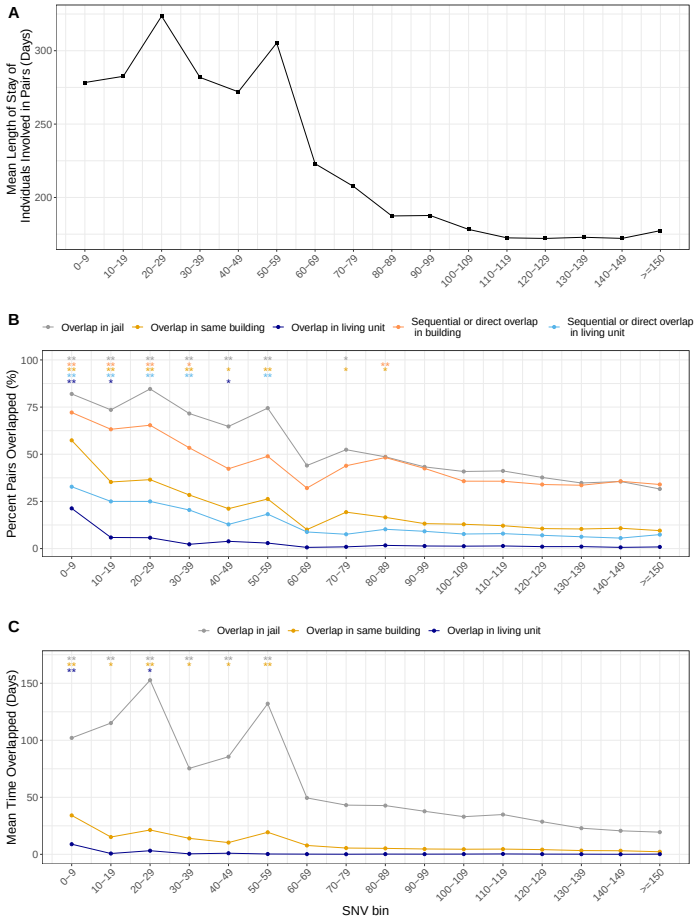
were within 20 SNVs of an intake isolate, 2 to another acquisition isolate, and 1 to a clinical isolate). We observed that 50.4% of jail-onset infections were within 20 SNVs of another isolate, with 17% of clinical isolates within 20 SNVs of an intake colonization strain. 95% of jail- onset clinical MRSA isolates with a genetic neighbor within 20 SNVs were related to another clinical MRSA strain. 76% of jail-onset clinical MRSA isolates that are within 20 SNVs of an- other isolate are within 20 SNVs of multiple isolates (range 2–8) (Supplementary Figure S10).

For detainees with USA300 isolates, individuals with longer lengths of stay tended to have a closer genetic pair (Figure 3A, Supplementary Figure S11). Furthermore, individuals who developed MRSA infection had longer jail lengths of stay than study participants who came into the jail MRSA colonized and did not develop infection during that incarceration (Supplementary Figure S12).

3.3.6 Relationship of jail location to USA300 MRSA colonization and clinical isolates

Eight buildings housed male detainees, including 4 cell-based and 4 dormitory-style buildings. Individuals incarcerated in the jail typically had opportunity for multiple movements, including between buildings (Supplementary Figure S13), to court, and to social programs (eg, church, school, drug treatment). Among individuals who acquired USA300 MRSA colonization or had a jail-onset USA300 MRSA infection, those with more closely related genetic neighbors tended to be in the jail at the same time and to overlap in particular buildings (Figure 3.3B, Supplementary Figure S14) and for significantly longer than did random pairs of individuals (Figure 3.3C, Supplementary Figures S15 and S16). Furthermore, 13 of 35 pairs of individuals who overlapped in the same building and whose MRSA isolates were within 9 SNVs overlapped at the more granular level of living unit, suggesting direct (eg,

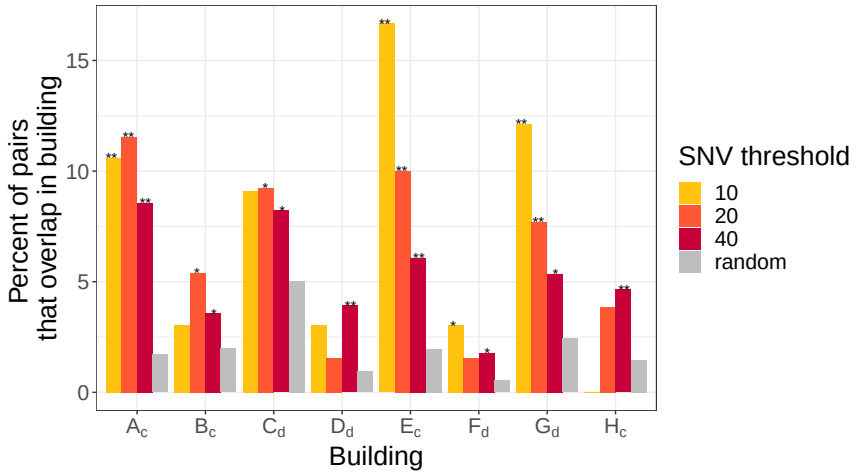
Figure 3.3: Individuals with closely related USA300 MRSA strains are more likely to reside in common jail locations and have longer length of stay



(A) Each square indicates the mean length of stay of unique individuals involved in a pair (y-axis) within the respective SNV distance range (x-axis). SNV distances ranges are inclusive. For A, B, and C, pairs involve a putative acquisition of USA300 MRSA (jail-acquired colonization or jail-onset infection) and a source. See distribution of length of stay in Supplementary Figure S11. (B) Each dot indicates the percentage of pairs related by the respective SNV distance that overlapped in the respective location (y-axis) in an epidemiologically relevant window. Sequential overlap indicates that 2 individuals were both in the same living unit at some point in their jail stay during an epidemiologically relevant window, but not necessarily at the same time. See Supplementary Figure 13 for results of permutation test. (C) Each dot indicates the mean time a pair of individuals related by the respective SNV distance overlapped in the jail, the same building, or the same living unit in an epidemiologically relevant window (see Supplementary Methods). See Supplementary Figure 14 to see the distribution of days overlapped in jail and in a particular building, and Supplementary Figure 15 for detailed results of permutation test. In panels B and C asterisks indicates significance by permutation test where 1 asterisk indicates significance at $P < .05$ and 2 asterisk indicates significance at $P < .005$.

no intermediaries) transmission (Figure 3.3B). While more distantly related pairs (10–50 SNVs) overlapped in the same building more than random pairs (>50 SNVs), they tended to have only sequential exposures to the same living units, suggesting potential transmission with intermediaries or environmental contamination as a source (Figure 3.3B, Supplementary Figure S14D, E).

Figure 3.4: Overlap in buildings among individuals with closely related MRSA strains



Colors represent pairs of USA300 isolates (including a putative acquisition of colonization or infection and a source) genetically related at different SNV thresholds (ie, $\leq 10, \leq 20, \leq 40$). Random (in gray) indicates all pairs of isolates and is shown to provide the baseline location sharing of pairs of individuals for each building regardless of genetic linkage. The x-axis indicates the different buildings male detainees could stay, labeled as cell-based (subscript c) or dorm-based building (subscript d). The y-axis indicates the percent of pairs that overlap in an epidemiologically relevant window in each building (See Supplementary Methods). Asterisks indicates significance by permutation test where 1 asterisk indicates significance at $P < .05$ and 2 asterisks indicates significance at $P < .005$ (see Supplementary Methods and Supplementary Figure S17 for results of permutation test).

While all buildings were sites of overlap among individuals harboring closely related MRSA strains, certain buildings, including both dorm and cell-based buildings, had significantly more overlap than expected by chance (Figure 3.4, Supplementary Figure S17). We also noted that most genomic clusters of individuals whose isolates are within 20 SNVs (range 2–8 isolates) cannot be explained by overlap in a single building, indicating that transmission

clusters are not necessarily confined to individual buildings (Supplementary Figure S10).

3.4 Discussion

There is a high burden of MRSA entering Cook County Jail; 19% of males arrive colonized [10]. Beyond that, we detected presumptive acquisition of MRSA colonization by day 30 of incarceration in an additional 8.4%, with sharing personal items a major risk factor. Genomic analyses, especially the small SNV differences among acquired and clinical MRSA isolates, suggests potential spread of incoming as well as of prevalent MRSA strains, with transmission potentially occurring among detainees housed together.

The 8% acquisition rate is higher than that reported among individuals in other congregate settings [24] and more in line with rates in intensive care units [25, 26] where MRSA, at least for some units, may be viewed as endemic. Interestingly, we ob- served that heroin use before entering the jail was significantly associated with acquisition. It is unclear if individuals who use heroin tend to congregate with certain populations, are housed in similar locations, or are characterized by factors or behaviors occurring during incarceration that were unmeasured. Nevertheless, a unique feature and likely a major challenge of jails is the exceedingly high incoming MRSA prevalence [10].

While it did not attain statistical significance, we observed that more frequent showering was negatively associated with acquisition of MRSA colonization. A prior case-control study in the LA County Jail observed that sharing soap with other inmates and less frequent showering were 2 factors associated with developing a MRSA infection during incarceration [89]. Individuals who developed MRSA infection tended to have longer lengths of detention, which may reflect the longer period of observation but also more opportunities for interactions with others. In the community, recommendations for hygiene and against sharing personal items remain critical components of education patients receive to prevent MRSA

[77]. Education regarding sharing of personal items and hygiene is a key, although difficult to enforce, MRSA control intervention for congregate living settings such as jails.

Using WGS, we observed that clinical USA300 infections occurring during incarceration have greater genomic similarity to each other in comparison to the diverse intake USA300 colonization strains. This finding suggests that infections could have originated from transmission within the jail. We observed that individuals with genetically similar MRSA strains were more likely to overlap in jail, suggesting spread among detainees housed in similar locations. As support for possible transmission, 35 of the 61 pairs within 9 SNVs overlapped in the same building, with 20 pairs having directly (n = 13) or sequentially overlapped in the same living units. While pairs separated by moderate SNV thresholds (20–50) still had significant overlap in a particular building, they often did not directly overlap in a living unit within that building. However, 10%–16% of pairs at moderate SNV thresholds had sequential occupancy of the same living unit, suggesting persistent environmental contamination or exposure to a more persistent MRSA strain in the jail.

Our observation that clusters of genomically-similar infection isolates existed even among people not sharing the same building or living unit suggests there could be virulent sublineages of MRSA that are more likely to cause infections or certain infections (eg, draining wounds) are more likely to be involved in MRSA spread. These hypotheses warrant additional investigation as they could provide possible targets for interventions. The continual influx of MRSA-colonized individuals into a setting characterized by close person-to-person contact, com- pounded by reduced opportunities for infection control, shows how challenging infection control can be in congregate settings. Prior mathematical models of the LA County Jail MRSA outbreak predicted that MRSA spread becomes more problematic as there are increased numbers entering the system with MRSA [7]; our results support this prediction. This model also noted that as more infections occurred during incarceration, increased spread to the surrounding community could occur [80]. Further understanding

both the downstream impact in the community of high MRSA burden in the jail as well as delineating the role recidivism contributes to overall MRSA burden is warranted.

Our study has limitations. First, we performed surveillance for MRSA colonization acquisition at Day 30 and thus may have missed acquisitions at shorter jail stays. Second, we cannot state definitively whether Day 30 positive swabs are acquisitions or intermittent colonization. However, the epidemiologic risk factors (ie, sharing personal items) and genomic data lend sup- port to these instances being acquisitions. Although given the small number of acquisitions, though statistically significant, the effects for heroin use and sharing personal items should be interpreted with due caution. Third, we did not ascertain infection control behaviors (eg, showering frequency, sharing personal items) among individuals who developed an MRSA infection and therefore cannot comment on the influence of such behavior on developing an infection. Fourth, we screened only a small percent of incoming detainees and likely missed some intake colonization that put detainees at risk for developing endogenous infections and serving as potential sources of MRSA transmission to others. Even with limited sampling, we observed that 17% of jail-onset clinical USA300 MRSA isolates were within 20 SNVs of an intake MRSA strain.

Our study demonstrates that not only is there a high level of MRSA colonization at jail entrance, acquisition of colonization and infection may occur during incarceration. Genomic analyses support this contention and suggest that spread occurred more frequently in certain jail locations. Sharing personal items was associated with acquisition of MRSA; an education campaign aimed at this practice could be a strategy to help curb spread. While our study examined MRSA, such an education campaign could even be extended to COVID-19, a pathogen that significantly impacted jails. Future study with more detailed epidemiologic analysis and environmental sampling within the jail complex might further inform and target interventions. Finally, the utility of an intervention at intake and/or discharge may be another focus of research as the jail remains a critical component of MRSA epidemiology in

urban areas.

Table 3.1: Epidemiological factors associated with acquisition of MRSA colonization

Epidemiologic Factor	MRSA Acquisition (n=12)[a]	No MRSA Acquisition (n=131)	OR	95% CI	P value
Exposures prior to incarceration					
Race/Ethnicity					
African-American	11 (92%)	106 (81%)			
(reference)					
Hispanic	0	10 (8%)	N/A	N/A	0.6
White/Other	1 (8%)	15 (11%)	0.64	0.07, 5.34	1.00
Age, mean years (SD)	39.7 (11.9)	37.6 (11.8)	1.02	0.97, 1.07	0.55
Heroin use in the past year	5 (42%)	21 (16%)	3.67	1.06, 12.68	0.05
Marijuana use in the past year	8 (67%)	89 (69%)	0.9	0.26, 3.16	1.00
Cocaine use in the past year	5 (42%)	48 (37%)	1.21	0.36, 4.01	0.76
Ecstasy or psychedelic use in the past year	1 (8%)	31 (24%)	0.29	0.04, 2.32	0.30
Other narcotics (e.g., codeine, oxycontin) in the past year	2 (17%)	14 (11%)	1.64	0.33, 8.27	0.63
Illicit benzodiazepine use in the past year	1 (8%)	13 (10%)	0.81	0.1, 6.8	1.00
Taking prescription drugs to get high in the past year	1 (8%)	6 (5%)	1.86	0.21, 16.9	0.47
Injection drug use in past year	2 (17%)	13 (10%)	1.82	0.36, 9.2	0.61
Homeless or unstable housing in the past year	7 (58%)	62 (47%)	1.56	0.47, 5.16	0.46
HIV infection	9 (75%)	91 (69%)	1.32	0.34, 5.13	1.00
Taking antiretrovirals	3 (33%)	59 (66%)	0.26	0.06, 1.12	0.08
Taking TMP-SMX	1 (8%)	13 (10%)	0.83	0.1, 6.91	1.00
Men who have sex with men	5 (42%)	34 (26%)	2.04	0.61, 6.85	0.31
ER visit in the past year	7 (58%)	71 (54%)	1.18	0.36, 3.92	0.78
Hospitalized in the past year	6 (50%)	50 (38%)	1.62	0.50, 5.30	0.54
Exposures during incarceration					
Participated in drug treatment	4 (33%)	22 (17%)	2.48	0.69, 8.95	0.23
Sharing of personal items[b]	7 (58%)	29 (22%)	4.92	1.45, 16.67	0.01
Any skin infections	1 (8%)	4 (3%)	2.89	0.30, 28.11	0.36
Visit to infirmary	4 (33%)	20 (15%)	2.77	0.76, 10.09	0.12
Number of times showered in the past week, mean (SD)	4.8 (1.8)	6.0 (2.7)	0.81	0.62, 1.07	0.13

[a] With inclusion of the new strain acquisition event detected with WGS for a person colonized at intake and remained colonized at DAY 30, all associations remained similar for predicting acquisitions. Heroin use before incarceration (P = .02) and sharing personal items (P = .02) both remained significant.
[b] Personal items shared by individuals who acquired MRSA included towel, toothpaste, uniform, and deodorant.

3.5 Methods

3.5.1 Study population

The study setting was the Cook County Jail in Chicago, IL, one of the largest single-site US jails, with roughly 250 incarcerations daily and daily census of 9000–10 000 detainees. Incarcerated males were enrolled within 72 h of entering jail from January 2016–December 2017. To enroll throughout the year and given the large number of HIV-negative individuals

entering the jail, we targeted enrollment at 10 HIV-negative males each week. Given our prior work demonstrating the significant impact MRSA has on HIV-infected individuals and the possible inter- section with incarceration for amplifying risk [39, 11, 90], we enriched the study population for HIV-infected individuals by enrolling from the jail HIV clinic (58% of our enrolled sample was HIV-infected). The estimated prevalence of HIV-infected detainees at the Cook County Jail is 2%. Individuals were enrolled from the jail HIV clinic within 24–48 h from jail entrance. Males were followed during incarceration and were eligible for a second study visit at Day 30 if still incarcerated.

3.5.2 Swab collection and processing

Surveillance cultures (anterior nares, throat, and bilateral inguinal) were collected at enrollment and Day 30 (if still incarcerated). Intake and Day 30 results determined colonization status. Specimens were obtained using the Copan ESwab for MRSA. Nasal swabs were collected by swabbing both anterior nares; throat swabs by swabbing the posterior pharynx; and inguinal swabs by swabbing a 10 cm2 skin area bilaterally [38]. Sample sites were chosen to maximize identification of MRSA carriers [76, 11]. Swabs were inoculated into enrichment broth to increase culture sensitivity [91]. Aliquots of overnight broth cultures were inoculated on ChromID MRSA (bioMérieux, North Carolina). MRSA was confirmed by standard biochemical tests; methicillin resistance by cefoxitin disk. Confirmed MRSA isolates underwent DNA extraction.

3.5.3 Archived clinical MRSA isolates

Existing archived clinical MRSA from male detainees incarcerated during the time of the study also underwent genomic sequencing. Clinical isolates from detainees who had been in the jail for > 72 h were considered "jail-onset" infections; those occurring ≤ 72 h into the

jail stay were considered community-onset. Infections from before the study, in an individual who remained in jail during the study, also were included to put into context acquisition MRSA isolates and to better characterize the genomic epidemiology of circulating strains (See Supplemental Methods).

3.5.4 Whole genome sequencing

Whole genome sequencing (WGS) was performed on MRSA isolates from jail entry and day 30 study visits and on clinical MRSA isolates from male detainees collected during the study period (See Supplementary Methods). Details on sequenced strains are available in Supplementary Table 1 and raw sequence data are available under Bioproject PRJNA638400. Intake positive colonization isolates were previously submitted under Bioproject PRJNA530184. Representative isolates that capture all independent acquisition events were selected for each person (See Supplementary Methods, Supplementary Figures S1 and S2) Publicly available USA300 genomes used in Figure 3.1 were downloaded from Bioproject PRJNA374377 [25] (Supplementary Table 2). Details on variant calling and phylogenetic analysis are in the Supplementary Methods.

3.5.5 Location overlap analysis

Transmission pairs were defined as involving one individual who acquired colonization or developed a jail-onset infection and a source. Potential sources were community-onset MRSA infections, infections that occurred outside the study period in an individual who remained in jail during the study, those already colonized at intake, other jail-onset MRSA infections, and other acquired colonization isolates (See Supplemental Methods, Supplementary Figure S3). Location sharing among genetically related transmission pairs was assessed for only USA300 isolates. Electronically available jail location data, including building and living

units, were ascertained for enrolled individuals and those with jail-onset infections. For each pair within the particular single-nucleotide variant (SNV) window (eg, 0–9, 10–19), overlap in the jail or in a particular building during an epidemiologically relevant window was calculated (Supplementary Figure S4). Based on the building in which each pair overlapped, we determined if the pair shared a living unit at the same or any time during their stay. Statistical significance of overlap at difference SNV thresholds was determined using permutation tests (See Supplementary Methods).

3.5.6 Risk factors and statistical analysis

A survey to identify predictors of MRSA colonization was ad- ministered to detainees at enrollment and included questions about drug use, sexual behaviors, housing status, and incarceration history. A survey about behaviors and activities during incarceration was administered at Day 30 to identify predictors of MRSA acquisition.

Using intake and Day 30 surveillance cultures, we deter- mined the frequency of persistent colonization (MRSA positive at intake and Day 30), presumptive acquisition (MRSA negative at intake and positive at Day 30), loss of colonization (MRSA positive at intake and negative at Day 30), and absence of colonization (negative at both time points). Results of surveillance cultures and intake surveys were used for risk factor analysis. SAS software version 9.4 (SAS Institute, Cary, North Carolina) was used for statistical analysis. Chi-square analysis was used for categorical variables, with Fisher's exact test for low-frequency predictors.

The study was approved by the Cook County Health institutional review board (IRB) which oversees approval for enrollment of jail detainees and the Rush University IRB; verbal consent was obtained. Approval from the Office for Human Research Protections was obtained to enroll current detainees.

3.6 Supplementary Data

Supplementary materials and figures are available at Clinical Infectious Diseases online at https://doi.org/10.1093/cid/ciaa1937.

3.7 Acknowledgements

We thank the Health Research and Solutions Center at Cook County Health for their assistance with data collection, Jon Zelner at the University of Michigan on guidance on statistical analyses, and Ali Pirani at the University of Michigan for bioinformatics support. We thank Bala Hota and Mary Hayden at Rush University Medical Center with their assistance with the early development of this study. We thank Connie Mennella, chair of correctional health at Cermak Health Services, with her assistance in the early planning of the project. We thank the individuals who participated in this study.

Chapter 4

Convergent Evolution of a Resistance-Conferring Plasmid in a Large, Urban Jail and the Broader Community

4.1 Preamble

Chapter 2 described MRSA epidemiology in the community of Cook County, and Chapter 3 identified transmission dynamics in the Cook County Jail. Chapter 4 explores the interplay of the jail and the community by uncovering a plasmid that is selected for in the jail, associated with transmission, and present in the broader community.

I performed the genomic, phylogenetic, and statistical analyses and created the figures presented in this chapter and drafted this chapter.

4.2 Introduction

Once a primarily healthcare-associated pathogen, MRSA has infiltrated the community to cause infections in otherwise healthy individuals [31]. Incarcerated individuals are one sector of the community with a high burden of MRSA colonization and infection. For example, we recently reported a striking MRSA colonization prevalence of 19% at the time of intake in the Cook County Jail (Chicago, IL) [38], much higher than the ~2% carriage prevalence in the general population across the U.S. [7]. In addition, jails are high transmission settings

for a variety of infectious diseases including MRSA [67], tuberculosis [55], and COVID-19 [92] as a result of close person-to-person contact.

The high transmission setting may promote selection of variants of pathogens [55]. Moreover, jails may impart different selective pressures than the healthcare setting or the broader community. As jails are characterized by short lengths of stay and high recidivism, there are multiple opportunities for acquisition in the jail and spread to the broader community. Thus, jails are hypothesized to be amplifiers of MRSA transmission in the community. Recently, it was shown that prisons could select for problematic variants (e.g. variants that are more transmissible or of increased antibiotic resistance) in *Mycobacterium tuberculosis* which were subsequently propagated in the community [55]. Thus, studying pathogen evolution in jails and prisons is important for infection control efforts in the jail and broader community.

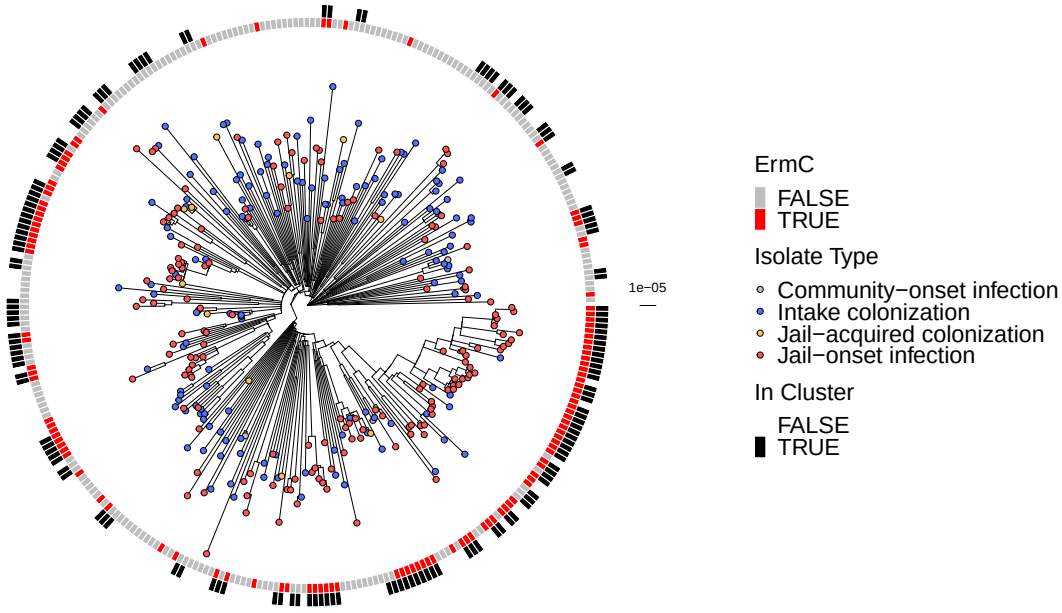
In this study, we sought to identify variants associated with USA300 MRSA transmission that were selected for in the Cook County Jail, where we previously observed evidence of MRSA transmission. We then looked for the presence of variants in the broader community in a comprehensive sample of clinical cultures to compare the strength of selection in the jail and community and the prevalence of the variants outside of the jail.

4.3 Results

4.3.1 Transmission among MRSA infections in the Cook County Jail

We previously conducted a genomic epidemiology analysis of USA300 MRSA transmission in the Cook County Jail [67]. Comparing genetic linkages among individuals entering the jail versus those acquiring colonization or infection in the jail revealed that jail-onset MRSA infections had closer genetic neighbors than MRSA imported into the jail from the community

Figure 4.1: Phylogenetic tree of USA300 MRSA in Cook County Jail



Maximum likelihood phylogeny of jail USA300 samples created with IQTREE. Scale bar indicates substitutions per site. Tips designate sample type. Inner ring indicates the presence (red) or absence (gray) of the *ermC*-carrying plasmid. Outer ring indicates if the isolate is genetically related to any other isolate within 20 SNVs (black).

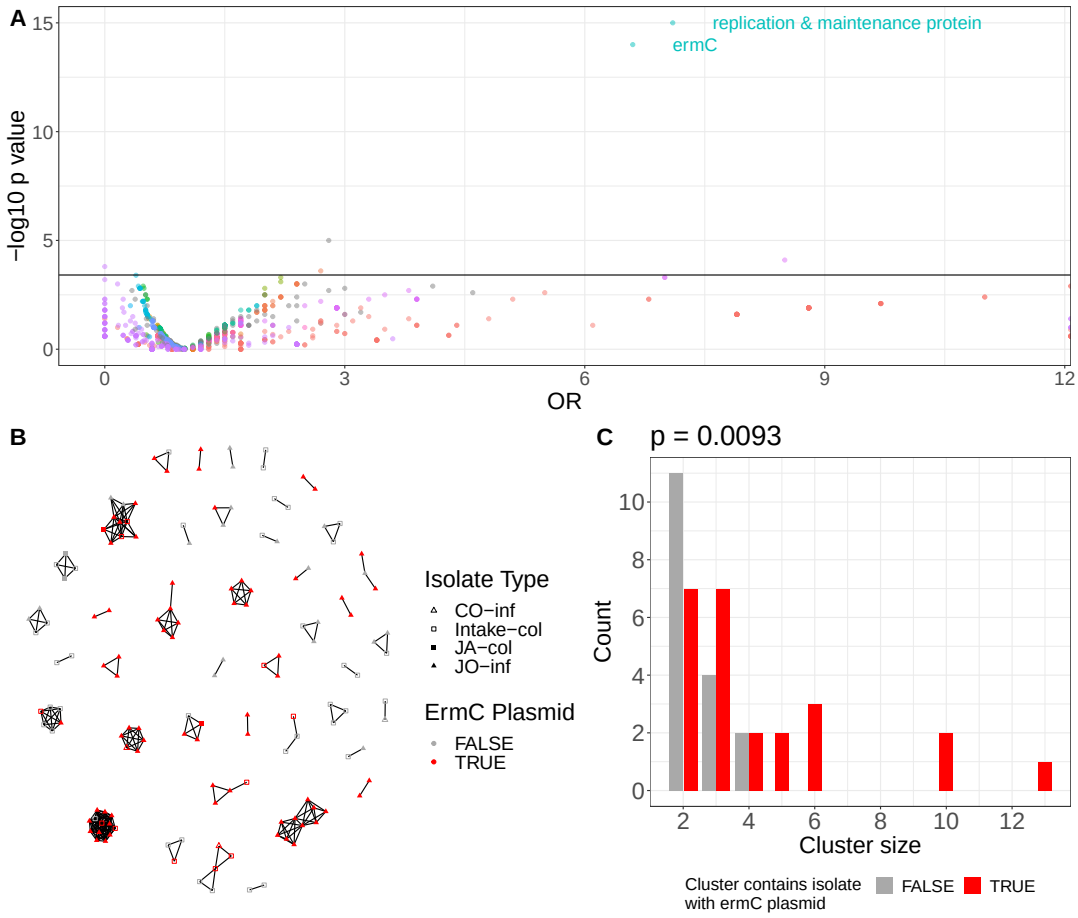
[67]. Moreover, examination of genetic linkages among jail-onset infections revealed that they form large clusters of individuals who harbored closely related strains and had spatiotemporal overlaps in the jail. In total, 157 USA300 jail-onset MRSA infections formed 25 clusters ranging in size from 2 to 10. These transmission clusters spanned the full diversity of USA300 (Figure 4.1). Moreover, 54.1% of jail-onset infections were genetically linked to another jail-onset infection indicating evidence of transmission among individuals with wound infections in the jail.

4.3.2 *ermC*-carrying plasmid associated with transmission in Cook County Jail from 2015-2017

A striking feature of jail-onset transmission linkages is their formation of large clusters, spread across the USA300 phylogeny. While there could be epidemiological explanations for this observation (e.g. super-spreader events or individuals), we wondered whether there were also microbial genetic contributors to the apparent elevated spread of certain sublineages. In particular, we wondered if there were shared, convergently evolved genetic underpinnings that promoted transmission. We focused on the accessory genome, as the gain and loss of genes and mobile genetic elements allows for rapid evolution [93]. We ran panaroo to determine accessory genes present in each genome and identified genes that may be traveling together on the same mobile genetic element by grouping genes with 95% concordance of presence-absence pattern across all isolates. This resulted in 37 gene clusters of size 2 to 211 (median size of 4) and 92 singletons.

We next assessed the association of genes and gene clusters with transmission, defined as being related to another isolate within 20 single nucleotide variants (SNVs). There were 6 genes found to be significantly associated with transmission after multiple test correction, with most being part of a gene cluster (Supplemental Table 4.2, Figure 4.2A). By several or-

Figure 4.2: Association between pangenome and transmission in the Cook County Jail from 2015-2017



A) Results of Fisher's exact tests for association between gene content and involvement in transmission cluster. Line indicates Bonferonni adjusted p value threshold. Colors indicate gene cluster, with gray indicating gene singletons. B) Network diagram of MRSA transmission clusters. Edges indicate that nodes are related within 20 SNVs. Nodes colored by presence or absence of the *ermC*-carrying plasmid. Shape indicates the isolate type, where triangles are infection and squares are colonization. C) Distribution of cluster size for clusters containing at least one isolate with the *ermC*-carrying plasmid and clusters with no isolates carrying the plasmid.

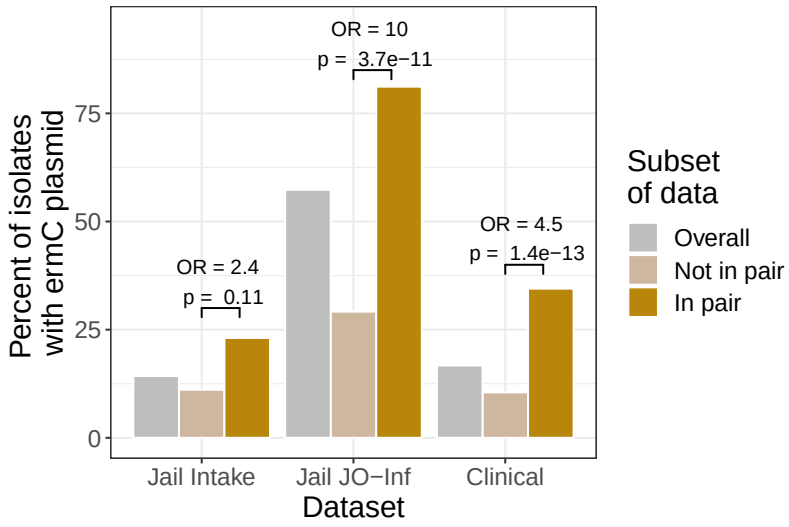
ders of magnitude, the most significant hits were *ermC* and a gene annotated as a replication and maintenance protein, which were the only 2 members of the inferred gene cluster. Comparison to sequence databases revealed that these genes are part of a small, 2.4 kb plasmid that confers resistance to macrolides and lincosamides and has been previously observed in USA300 MRSA [94, 95]. Acquisition of the *ermC*-carrying plasmid occurred across diverse genetic backgrounds, indicating that this plasmid was acquired multiple times by isolates of differing genetic backgrounds (Figure 4.1).

Not only is the *ermC*-carrying plasmid associated with being genetically linked to another isolate, individuals harboring MRSA with the *ermC*-carrying plasmid were part of significantly larger transmission clusters (Figure 4.2B,C) further implicating the role of the plasmid in proliferation within the jail. Indeed, the prevalence of isolates carrying this plasmid was higher among jail-onset infections than among isolates present at intake to the jail (57.3% vs. 14.3% respectively, Figure 4.3). This suggests that there were multiple importations into the jail, and that isolates containing this plasmid preferentially spread within the jail compared to isolates that do not. Despite the lower prevalence in the community, intake positive isolates with the *ermC*-carrying plasmid were still more often involved in transmission in the community, though not significantly so, potentially reflective of the role of the plasmid in transmission in the broader community (Figure 4.3).

4.3.3 *ermC*-carrying plasmid present in the larger community

To further explore the role of the *ermC*-carrying plasmid in community transmission, we conducted similar analyses in a comprehensive collection of clinical cultures from Cook County from 2011-2014. This dataset was collected in the years prior to the start of the jail dataset in 2015, therefore we could not directly assess the downstream effect of the jail on the community. The prevalence of this plasmid among the comprehensive clinical cultures was similar to the prevalence among MRSA colonization at intake to the jail (Figure 4.3).

Figure 4.3: Prevalence of *ermC*-carrying plasmid among datasets

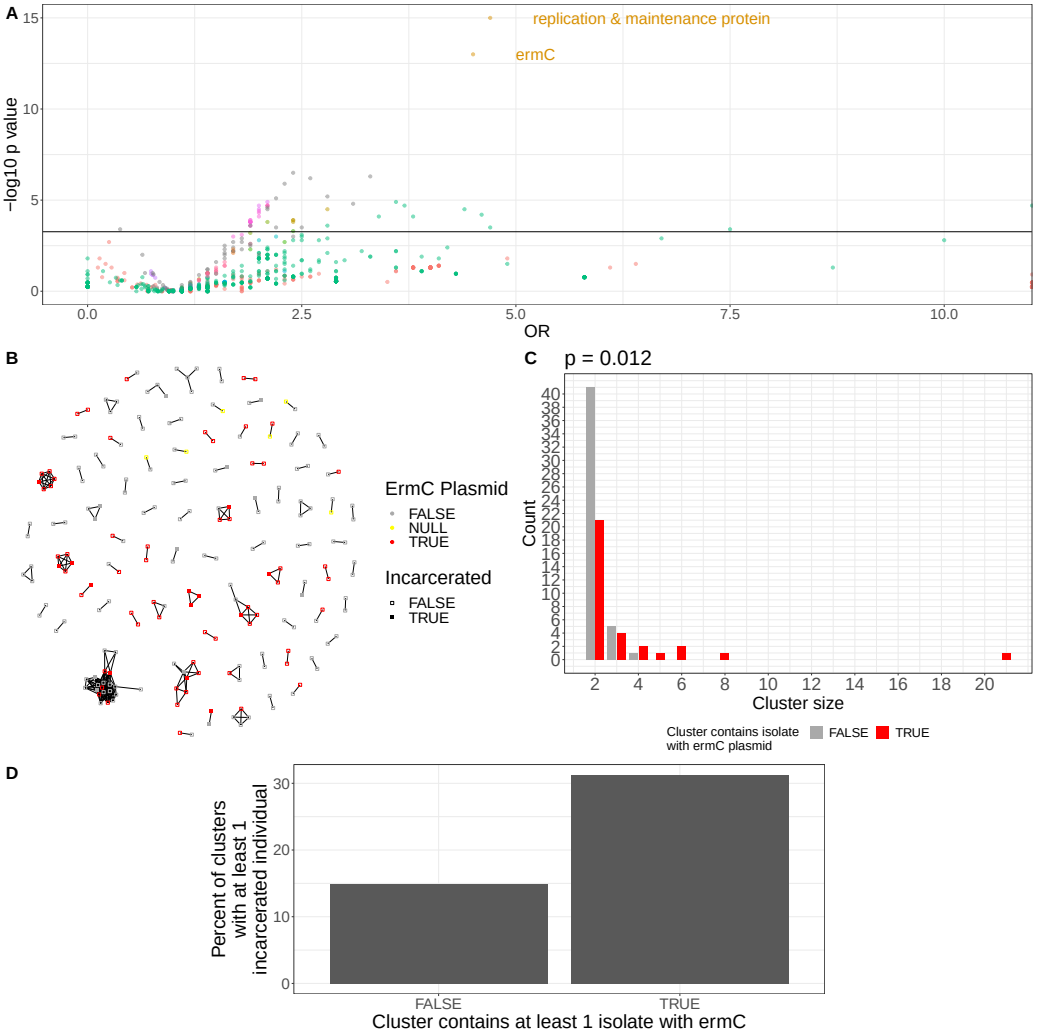


Gray bars indicate the overall prevalence of *ermC*-carrying plasmid among jail intake, jail-onset infection and comprehensive clinical cultures. Shades of yellow break indicate the prevalence of the *ermC*-carrying plasmid among isolates genetically linked to another isolate by 20 SNVs or not. Results of Fisher's exact test indicated.

We found that the same plasmid circulating in the jail was also most significantly associated with transmission (i.e. genetic linkage to another isolate by 20 SNVs) in the comprehensive clinical cultures from Cook County Health (Figure 4.4A, Supplemental Table 4.3). Again, this was a result of multiple acquisitions of the plasmid among isolates spanning the diversity of the USA300 tree (Supplemental Figure 4.7). Consistent with our findings in the jail, larger transmission clusters tend to contain isolates with the *ermC*-carrying plasmid (Figure 4.4B, C). Furthermore, the presence of the *ermC*-carrying plasmid is increasing over time, further supporting amplification in the larger community (Figure 4.5A) .

More genes were significantly associated with transmission in the clinical dataset than in the jail dataset (Figure 4.2A, 4.4A); these hits require further exploration (Supplemental Table 4.3). However, only one gene other than the *ermC*-carrying plasmid, annotated to contain a *Staphylococcal* superantigen-like OB-fold domain, is significant in both datasets but is low in prevalence in the jail (6.9%).

Figure 4.4: Association between *ermC* and transmission in comprehensive clinical cultures from 2011-2014 presenting to Cook County Health



A) Results of Fisher's exact tests for association between gene content and involvement in transmission cluster. Line indicates Bonferonni adjusted p value threshold. Colors indicate gene cluster, with gray indicating gene singletons. B) Network diagram of MRSA transmission clusters. Edges indicate that nodes are related within 20 SNVs. Nodes colored by presence or absence of the *ermC*-carrying plasmid. Shape indicates incarceration in the past year. C) Distribution of cluster size for clusters containing at least one isolate with the *ermC*-carrying plasmid and clusters with no isolates carrying the plasmid. D) Plot indicating the percent of clusters with at least 1 incarcerated individual among clusters where at least one isolate carries the *ermC* plasmid versus not.

4.3.4 *ermC*-carrying plasmid confers constitutive resistance to clindamycin

The *ermC*-carrying plasmid has been reported to confer resistance to macrolides and can confer constitutive or inducible resistance to the lincosamide clindamycin [94]. Constitutive resistance to clindamycin occurs through modifications in *ermC*or the leader peptide [94]. In the comprehensive dataset of clinical cultures among isolates carrying *ermC*, 110 had constitutive resistance to clindamycin, while 23 had inducible resistance. Among jail infections with resistance data that carried *ermC*, 56 had constitutive resistance to clindamycin, while 5 had inducible resistance.

We confirmed that the *ermC*-carrying plasmid was the most associated with clindamycin resistance using pyseer (Supplemental Figure 4.8), and phenotypically we observe constitutive resistance. However, we have not identified a SNV or indel associated with constitutive resistance in these particular plasmids. Future directions will look for large insertions that may confer constitutive resistance.

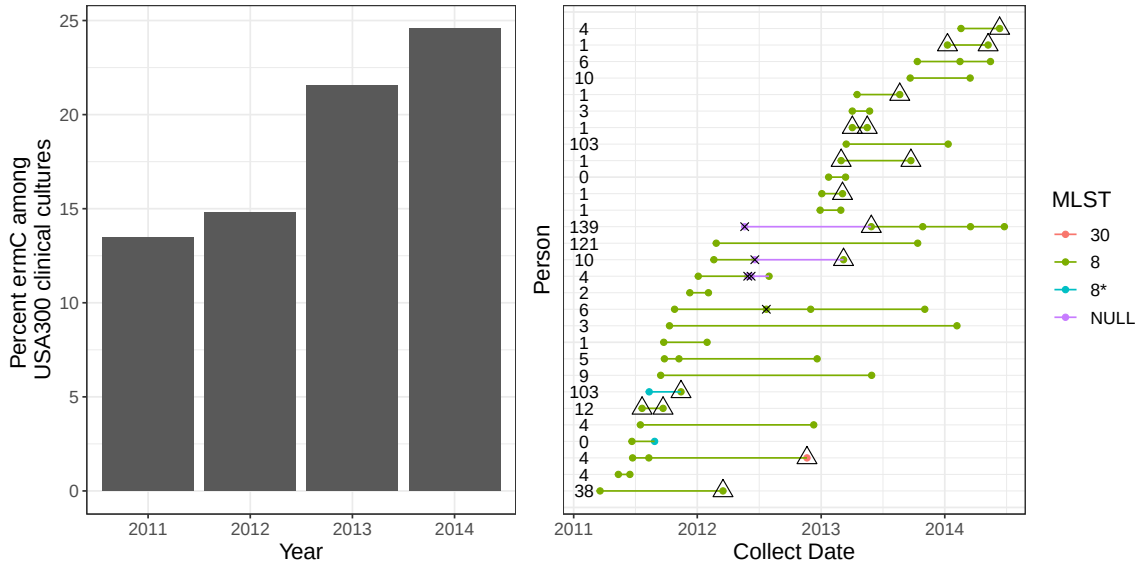
4.3.5 Epidemiological associations with *ermC*-carrying plasmid

Next we assessed if there were epidemiological characteristics of individuals harboring MRSA with the *ermC*-carrying plasmid in the comprehensive clinical dataset. Incarceration in the past year of admit date is associated with the presence of the *ermC*-carrying plasmid in the clinical dataset, suggesting it is possible that some strains were acquired in the jail and spread in the community. Indeed, we do see that clusters that contain at least 1 individual with *ermC*have a higher percentage of incarceration (Figure 4.4D, $p = 0.10$ OR = 2.60). Current use of cocaine is also significantly associated with carrying the *ermC*plasmid. Cocaine use is associated with incarceration in the prior year (OR = 2.5, $p = 0.0048$) and thus the association with the plasmid could be either a proxy for incarceration or reflective of

Table 4.1: Epidemiological associations with harboring the *ermC*-carrying plasmid among the comprehensive sampling of clinical cultures from 2011-2014

	Percent Epi Factor Among Those with *ermC*	**Percent Epi Factor Among Those without *ermC***	**p**	**OR**	**95% CI**
Trimethoprim-Sulfamethoxazole exposure in past 6 mo	25% (33/131)	13% (88/653)	0.0013	2.2	1.4,3.4
Vancomycin exposure in past 6 mo	43% (56/131)	29% (192/653)	0.0038	1.8	1.2,2.7
MRSA in past 6 mo	18% (23/131)	9% (59/653)	0.0071	2.1	1.2,3.6
Current cocaine user	21% (25/117)	12% (77/620)	0.013	1.9	1.2,3.2
Incarcerated in prior year	15% (19/131)	7.5% (49/653)	0.016	2.1	1.2,3.7
Pediatric patient (<13 years of age)	1.5% (2/131)	6.1% (40/653)	0.032	0.24	0.04,0.89
Cocaine use ever	31% (37/119)	22% (138/623)	0.045	1.6	1,2.5
Inpatient in prior year	31% (40/131)	24% (154/653)	0.097	1.4	0.93,2.2
Currently homeless	12% (15/129)	7% (45/644)	0.1	1.7	0.92,3.3
Wound infection (vs. other)	90% (118/131)	85% (552/653)	0.1	1.7	0.91,3.1
Current illicit drug user	42% (50/118)	35% (215/620)	0.12	1.4	0.92,2.1
Asthma	21% (28/131)	16% (103/653)	0.12	1.5	0.91,2.3
Clindamycin exposure in past 6 mo	62% (81/131)	54% (354/653)	0.12	1.4	0.92,2
History of illicit drug use	57% (68/119)	49% (307/623)	0.13	1.4	0.92,2
History of injection drug use	14% (17/119)	9.3% (58/623)	0.13	1.6	0.88,2.9
Ceftriaxone exposure in past 6 mo	6.9% (9/131)	11% (74/653)	0.16	0.58	0.28,1.2
Azithromycin exposure in past 6 mo	11% (15/131)	7.7% (50/653)	0.16	1.6	0.83,2.9
HIV	11% (15/131)	7.8% (51/653)	0.17	1.5	0.81,2.8
COPD	9.9% (13/131)	6.4% (42/653)	0.19	1.6	0.83,3.2
Illicit drug use in past 3 mo	23% (27/118)	18% (110/622)	0.2	1.4	0.85,2.3
Ciprofloxacin exposure in past 6 mo	8.4% (11/131)	5.7% (37/653)	0.23	1.5	0.75,3.1
Hospitalization in prior year	28% (37/131)	24% (154/653)	0.27	1.3	0.83,2
Hypertension	41% (54/131)	36% (234/653)	0.27	1.3	0.85,1.9
Levofloxacin exposure in past 6 mo	9.9% (13/131)	7% (46/653)	0.28	1.5	0.76,2.8
Current marijuana use	24% (28/117)	20% (124/620)	0.32	1.3	0.79,2
History of heroin use	24% (28/119)	19% (120/623)	0.32	1.3	0.81,2.1
Emergency room in prior year	44% (57/131)	39% (253/653)	0.33	1.2	0.83,1.8
History of incarceration	21% (28/131)	18% (119/653)	0.39	1.2	0.76,1.9
HCV	6.9% (9/131)	5.1% (33/653)	0.4	1.4	0.64,3
Current injection drug user	8.5% (10/117)	6.4% (40/622)	0.42	1.4	0.62,2.9
Diabetes	27% (35/131)	23% (152/653)	0.43	1.2	0.78,1.8
Current heroin user	17% (20/117)	15% (91/620)	0.48	1.2	0.69,2
History of marijuana use	36% (43/119)	33% (204/623)	0.52	1.2	0.77,1.8
Penicillin exposure in past 6 mo	1.5% (2/131)	2.8% (18/653)	0.55	0.55	0.09,2.2
Cefazolin exposure in past 6 mo	3.8% (5/131)	5.1% (33/653)	0.66	0.75	0.27,2
Black / African-American race (vs. other)	63% (82/131)	60% (394/653)	0.7	1.1	0.74,1.6
Cirrhosis	0.76% (1/131)	1.8% (12/653)	0.71	0.41	0.019,2.6
Outpatient in prior year	47% (62/131)	46% (299/653)	0.77	1.1	0.72,1.6
Cancer	11% (14/131)	12% (78/653)	0.77	0.88	0.46,1.6
ICU encounter in past 2 weeks	2.3% (3/131)	3.1% (20/653)	0.78	0.74	0.18,2.5
Surgery in prior year	17% (22/131)	18% (115/653)	0.9	0.94	0.56,1.6
Hispanic	21% (27/129)	22% (142/651)	0.91	0.95	0.59,1.5
ICU encounter 2 weeks after MRSA	1.5% (2/131)	1.5% (10/653)	1	1	0.15,4.3
Stroke	2.3% (3/131)	2.1% (14/653)	1	1.1	0.26,3.6
Augmentin exposure in past 6 months	6.1% (8/131)	6.4% (42/653)	1	0.95	0.4,2

Figure 4.5: Increasing prevalence and evidence of acquisition of *ermC*-carrying plasmid among repeat infections



A) Barplot shows the increasing prevalence of *ermC* from 2011 to 2014 among comprehensive clinical cultures. B) Evidence of *ermC* gained over time among individuals with repeat infections. Each y axis tick indicates a person with at least 1 repeat USA300 infection. Triangles indicate presence of *ermC*. Color indicates MLST. Numbers along y axis indicate max number of SNVs between USA300 infections; note this excludes any other MLST (i.e. the pink ST30 isolate is not included in the distance calculation).

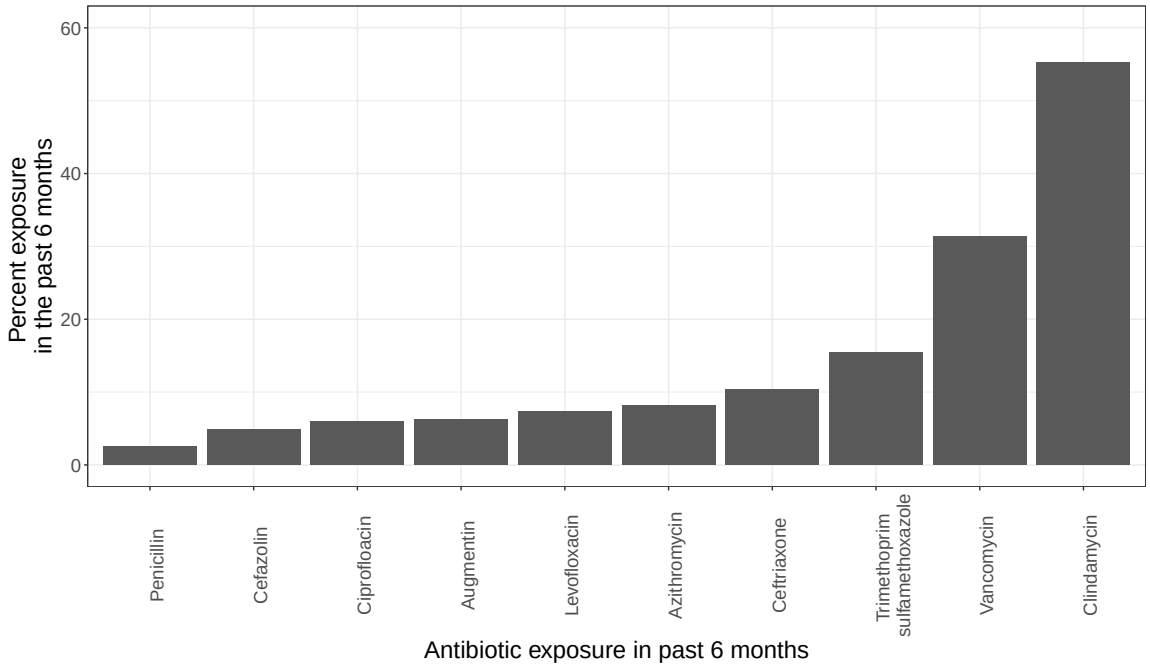
either a selective pressure or circulation in certain social networks. The pediatric population is negatively associated with carrying the *ermC* plasmid, suggesting circulation in adult populations as reflected in the association with incarceration and cocaine use.

Individuals could have repeat infections in the dataset, but we included individuals just once, selecting their first USA300 infection. Prior MRSA infection or colonization in the past 6 months and two antibiotics commonly used to treat MRSA (i.e. trimethoprim-sulfamethoxazole and vancomycin) were associated with *ermC* (Table 1). This could indicate that acquisition of the plasmid was recent wherein the first infection did not contain the plasmid but was present in the second infection. To test this directly, we identified individuals with at least one repeat USA300 infection, and saw evidence of *ermC* acquisition

in later infections of the same strain within 20 SNVs among 4 individuals (Figure 4.5B).

Naturally, we hypothesized that acquisition of the *ermC*-carrying plasmid was selected for by clindamycin or macrolide exposure. Interestingly, clindamycin and macrolide exposure in the past 6 months are not associated with presence of *ermC*(Table 4.2). However, 60% of individuals have been prescribed clindamycin in the past 6 months (Figure 4.6) which could explain the lack of association given that clindamycin is a ubiquitous exposure in the dataset. Further, we wondered if clindamycin exposure was particularly prevalent in the jail, given the increased selection for the plasmid in the jail. We did not have data on clindamycin use in the jail study period of 2015-2017. However, we assessed clindamycin use in the past 6 months and incarceration in the past year in the comprehensive clinical dataset collected from 2011-2014 and found no significant association (OR = 1.09, p = 0.80).

Figure 4.6: Antibiotic exposure in the past 6 months



Percent of individuals with a USA300 MRSA infection that were exposed to each antibiotic in the past 6 months among clinical data

4.4 Discussion

We identified a plasmid associated with transmission that confers constitutive resistance to clindamycin which has independently evolved multiple times in diverse genetic backgrounds and spread in both the Cook County Jail and larger community. Prevalence is much higher in the jail than in the broader community suggesting a stronger selective pressure in the jail. However, increasing prevalence in the community indicates a fitness advantage in the community.

Around this time, prevalence of clindamycin resistance in MRSA in the United States was reported to be on the order of 10%, but prevalence varies greatly by population and geography [96, 97, 98]. In the jail, we see prevalence of clindamycin resistance of over 50%. Further, there are many strains in the jail that are resistant to clindamycin and fluoroquinolones, a pattern atypical of USA300 [99], suggesting different selective pressures in the jail than the broader community. Prevalence of clindamycin resistance in the broader community prior to the jail study time period were on par with reports, but the association of transmission with presence of the *ermC*-carrying plasmid is novel.

The selective pressure for this phenomenon is unclear, as we do not observe a significant association between clindamycin or macrolide use and the presence of this plasmid. Previous studies revealed a plasmid pUSA03 which carries *ermC*and mupA has emerged in the MSM community in Boston and San Francisco [54]. Here too clindamycin and mupirocin use was not a necessary condition for the spread of this clone of USA300, as these drugs were not frequently used in one clinic where they observed this phenomenon [54]. Despite the lack of association, clindamycin was the most frequently prescribed antibiotic in the MRSA infected population presenting to Cook County Health. The ubiquitous selective pressure of clindamycin is constituent with the high prevalence of constitutive resistance. We are lacking data on clindamycin use in the Cook County Jail, but incarceration in the past year

was not associated with clindamycin exposure in the past 6 months in the comprehensive clinical dataset. Although the selective pressure for the *ermC*-carrying plasmid is unknown, it does seem to be higher in the jail than in the community.

In addition to antibiotics, there may be other selective pressures that favor isolates harboring the *ermC*plasmid. One intriguing selective agent is the additive triclosan, which is found in many consumer products including personal hygiene products (e.g. toothpaste, deodorant, mouthwash, soup), textiles, and plastics for its antibacterial properties [100, 101]. Studies have shown that triclosan in the environment is associated with higher levels of *erm* genes in the environment [102] and that triclosan may promote antibiotic resistance and tolerance [103, 104]. Further, triclosan has been shown to promote nasal colonization of *S. aureus* [105]. Among a representative sample of the U.S. population via the National Health and Nutrition Examination Survey, triclosan was found in the urine of 75% of individuals in the study, with more than 10% of the individuals having concentrations greater than the MIC for S. aureus [100, 104]. While triclosan exposure is ubiquitous, it is unclear if this selective pressure is present or stronger than average in the jail.

One important and unanswered question is the impact of jails on propagating the community MRSA epidemic. In this study, our clinical samples (2011-2014) were collected prior to the jail samples (2015-2017), limiting our ability to directly study the impact of the jail on the downstream community. Sequencing the archived clinical isolates from Cook County Health in 2015 and beyond and observing the intersection between isolates in the jail and the broader community could help quantify the contributions of the jail to the burden of MRSA in the downstream community. With the data we have, we do observe a significant association between presence of the *ermC*-carrying plasmid and previous incarceration. Further work is needed to understand the interplay between the jail and the community in the spread of the plasmid.

4.5 Methods

4.5.1 Data - USA300 MRSA whole-genome sequences in the Cook County Jail and Cook County Health

Metadata and MRSA sequences were previously described in Chapters 2 and 3. Briefly, all clinical cultures presenting to Cook County Health from 2011-2014 were sequenced (Chapter 2). We subsetted the dataset to USA300 MRSA, the dominant molecular type. MRSA surveillance colonization samples collected as part of prospective epidemiology study and all infection samples at the Cook County Jail were sequenced from 2015-2017 (Chapter 3).

4.5.2 Assessing association of transmission with pangenome

We used panaroo to identify the accessory genome of each sample [106]. Clusters of genes as a proxy for genes potentially carried on the same mobile genetic element were defined by concordance of presence-absence pattern across samples by 95%. Genes were annotated with eggNOG [107]. Isolates related by transmission were defined with a SNV threshold of 20 to minimize false positive linkages [71, 70]. A Fisher's exact test was conducted to assess the association between transmission and gene using the R package exact2x2. The relationship between odds ratio (OR) and p value was plotted and colored by gene cluster. Significance was assessed with a Bonferonni-adjusted p value where the number of tests was the number of clusters plus the number of singleton genes.

4.5.3 Genetic determinants of clindamycin resistance

*ermC*can confer constitutive or inducible resistance to clindamycin. In a clinical setting, constitutive resistance is observed phenotypically as resistance to clindamycin in the absence of an inducer (i.e. erythromycin). The potential for inducible resistance to clindamycin

through carriage of an unmodified erm gene would be investigated if there was resistance to erythromycin but susceptibility to clindamycin [108]. We used pyseer [109] to confirm that the *ermC*-carrying plasmid was the sole determinant of clindamycin resistance.

4.5.4 Identifying and comparing plasmid between datasets

We assembled the genomes with an internal pipeline available at https://github.com/alipirani88/assemblage. We then used blastn [110] to compare the assembly against the plasmid (with inducible *ermC*) pUSA05-1-SUR4 (NCBI reference NZ_CP014374.1) [94]. We also used blastn to confirm the identity of the 2 highly significant gene hits present in both datasets.

4.5.5 Epidemiological associations with *ermC*-carrying plasmid

We assessed the association between binary epidemiological factors and presence of the *ermC*-carrying plasmid with a Fisher's exact testing using the R package exact2x2.

4.6 Supplemental Tables

Table 4.2: Pangenome genes that pass the Bonferonni correction for jail dataset

gene	cluster	labels	pvals	OR	CI	log10p	eggNOG identifier	Preferred name	eggNOG description	prevalence
group_2305	26	replication & maintenance protein	3.1E-15	7.1	4.2,12	15	NA	NA	NA	36%
ermC	26	ermC	2.2E-14	6.6	4,11	14	NA	NA	NA	36%
group_467	NA		1.1E-05	2.8	1.7,4.4	5	1280.SAXN108_2730		Protein of unknown function (DUF1433)	52%
group_398	36		7.3E-05	8.5	2.5,34	4.1	1280.SAXN108_0481		Staphylococcal superantigen-like OB-fold domain	6.9%
group_497	35		0.00017	0	0,0.3	3.8	525378.HMPREF0793_1693	blaR	Regulatory protein BlaR1	4.4%
patA_2~~~patA_3	10		0.00024	2.7	1.5,4.7	3.6	NA	NA	NA	73%

Table 4.3: Pangenome genes that pass the Bonferonni correction for comprehensive clinical culture dataset

gene	cluster	labels	pvals	OR	CI	log10p	eggNOG identifier	Preferrenname	eggNOG description	prevalence
group_3371	11	replication & maintenance protein	6.4E-16	4.7	3.2,6.9	15	NA	NA	NA	19%
ermC~~~\rmC	11	\ermC	1.4E-13	4.5	3.6,7	13	NA	NA	NA	17%
group_1270	NA		3E-07	2.4	1.7,3.3	6.5	1280.SAXN108_1829			33%
lytN_3~~~lytN_4~~~lss~~~lytN_2	NA		5.4E-07	3.3	2.1,5.4	6.3	NA	NA	NA	11%
group_976	NA		6.4E-07	2.6	1.8,3.8	6.2	525378.HMPREF0793_0195	tcmP	Leucine carboxyl methyltransferase	21%
group_995	NA		1.2E-06	2.3	1.6,3.2	5.9	1280.SAXN108_0496		multivesicular body membrane disassembly	34%
group_425	NA		6.1E-06	2.8	1.8,4.4	5.2	1280.SAXN108_1511		this gene contains a nucleotide ambiguity which may be the result of a sequencing error	12%
group_695	NA		7.1E-06	2.2	1.6,3.1	5.1	1280.SAXN108_0973		Major facilitator Superfamily	28%
group_1128	6		1.4E-05	2.1	1.5,2.9	4.9	1220551.SCHR_11194	penP	beta-lactamase	41%
group_599	17		1.4E-05	3.6	2.6,3	4.9	1280.SAXN108_1250	femA	protein involved in methicillin resistance	6.8%
group_1442	NA		1.7E-05	3.1	1.9,5.2	4.8	1280.SAXN108_2114	fhuD	Periplasmic binding protein	8.5%
group_613	6		2.1E-05	2	1.5,2.8	4.7	1280.SAXN108_0734		AraC family transcriptional regulator	40%
group_339	8		2.2E-05	2.1	1.5,3	4.7	1280.SAXN108_1837		Nacht domain	27%
group_840	8		2.2E-05	2.1	1.5,3	4.7	71366.O80066_9CAUD		cytolysis in other organism	27%
group_242	8		2.2E-05	2.1	1.5,3	4.7	1280.SAXN108_0656		Protein of unknown function (DUF443)	27%
group_1130	17		2.2E-05	3.7	2.1,7	4.7	904314.SEVCU012_1101	penP	beta-lactamase	6.1%
group_1363	17		1.8E-05	Inf	5.1,Inf	4.7	176280.SE_0037		Recombinase zinc beta ribbon domain	1%
group_1358	NA		2.6E-05	2.1	1.5,2.9	4.6	NA	NA	NA	32%
atl_4~~~atl_3~~~atl_2~~~atl_1	6		3.1E-05	2	1.5,2.8	4.5	NA	NA	NA	41%
group_986	NA		3.1E-05	2.1	1.5,3.1	4.5	186152.Q8SDS6_9CAUD		Phage integrase family	25%
group_1320	17		3.3E-05	4.4	2.2,9	4.5	NA	NA	NA	4.3%
group_273	12		3E-05	2.8	1.7,4.6	4.5	1280.SAXN108_1513		Pfam PF07901	9.8%
group_1074	7		3.9E-05	2	1.4,2.8	4.4	1280.SAXN108_0656		Protein of unknown function (DUF443)	35%
group_651	7		5.3E-05	2	1.4,2.8	4.3	1140002.I570_02054			35%
group_1325	17		5.8E-05	4.6	2.1,9.9	4.2	NA	NA	NA	3.8%
group_1140	7		7.5E-05	2	1.4,2.8	4.1	176279.SERP2453		Belongs to the staphylococcal tandem lipoprotein family	35%
group_1561	17		8.3E-05	3.8	1.9,7.5	4.1	NA	NA	NA	4.8%
group_1721	17		8.8E-05	3.4	1.8,6.5	4.1	1280.SAXN108_1856		cell killing	5.9%
yqbO_1 ~~~ yqbO_2 ~~~~~~~ yqbO ~~~ yqbO_3	7		0.00012	1.9	1.4,2.7	3.9	NA	NA	NA	35%
group_921	7		0.00012	1.9	1.4,2.7	3.9	1280.SAXN108_2736		helicase	35%
group_1045	12		0.00012	2.4	1.5,3.7	3.9	NA	NA	NA	13%
group_1641	12		0.00012	2.4	1.5,3.7	3.9	NA	NA	NA	13%
group_1307	7		0.00017	1.9	1.3,2.6	3.8	NA	NA	NA	36%
group_1339	7		0.00017	1.9	1.3,2.6	3.8	NA	NA	NA	35%
group_1181	7		0.00017	1.9	1.3,2.6	3.8	435838.HMPREF0786_01635		Domain of unknown function (DUF927)	36%
group_1004	7		0.00017	1.9	1.3,2.6	3.8	129009.Q9MBT7_9CAUD			35%
group_1360	7		0.00017	1.9	1.3,2.6	3.8	NA	NA	NA	35%
group_1319	7		0.00017	1.9	1.3,2.6	3.8	1280.SAXN108_2879	yhgC	enzyme involved in biosynthesis of extracellular polysaccharides	35%
group_589	7		0.00016	1.9	1.3,2.6	3.8	1177179.A11A3_17040		reverse transcriptase	35%
group_400	14		0.00015	2.1	1.4,3	3.8	1280.SAXN108_0481		Staphylococcal superantigen-like OB-fold domain	22%
xis	12		0.00016	2.4	1.5,3.8	3.8	106284.Q9G032_9CAUD			12%
group_1354	7		0.00024	1.9	1.3,2.6	3.6	1280.SAXN108_0652		LXG domain of WXG superfamily	35%
group_1213	7		0.00024	1.9	1.3,2.6	3.6	1139219.I569_00564		Psort location CytoplasmicMembrane, score	35%
xerC_4~~~xerC_2~~~intQ	NA		0.00025	1.9	1.3,2.6	3.6	NA	NA	NA	30%
group_482	17		0.00023	2.8	1.6,4.8	3.6	525378.HMPREF0793_1527		Domain of unknown function (DUF5079)	7.9%
group_1208	NA		0.00029	2.2	1.5,3.4	3.5	1167632.AJTR01000009_gene931	binR	Resolvase, N terminal domain	15%
group_282	NA		3E-04	2.5	1.5,4.1	3.5	1280.SAXN108_1507		Pfam PF07901	9.8%
repD~~~repE_2~~~repN~~~repN_2	17		0.00034	4.7	2,12	3.5	NA	NA	NA	2.9%
group_1291	NA		0.00041	0.38	0.22,0.64	3.4	1280.SAXN108_1936		DNA packaging	92%
group_1421	17		0.00041	7.5	2.2,26	3.4	1280.SAXN108_1450	ypgR	virulence factor	1.8%
group_1324	15		0.00047	2.4	1.5,3.8	3.3	NA	NA	NA	10%

4.7 Supplemental Figures

Figure 4.7: Combined phylogeny of clinical and jail USA300 MRSA cultures

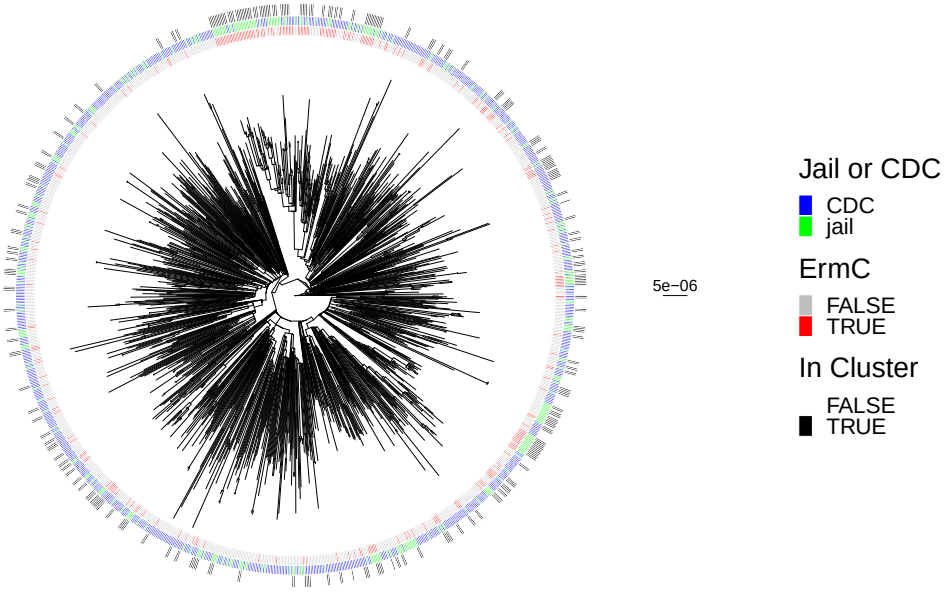
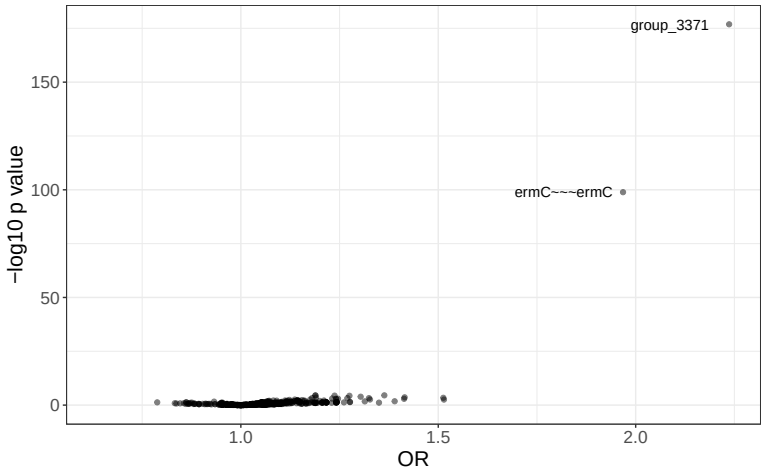


Figure 4.8: Pyseer results assessing association between clindamycin resistance and the pangenome in the clinical data



Chapter 5

Discussion

5.1 Major dissertation contributions

This dissertation produced insights into USA300 MRSA epidemiology and adaptation while demonstrating the utility of genomic epidemiology in both the healthcare and community settings. In Chapter 2, we described the genomic epidemiology of USA300 MRSA in Cook County, which contains the city of Chicago, from 2011-2014. We found little evidence of healthcare overlap among individuals with genetically-linked MRSA, even among so-called healthcare-associated or hospital-onset infections, thus calling into question the utility of these oft-used epidemiological definitions. In Chapter 3, we focused on one section of the community with a higher risk of MRSA colonization and infection: incarcerated individuals. We observed a high burden of MRSA at intake to the jail, followed by evidence of transmission mediated by location sharing and environmental contamination. Finally, in Chapter 4, we identified a biological underpinning of transmission in the Cook County Jail: acquisition of a 2.4kb plasmid carrying the antibiotic resistance determinant *ermC*. We found that this plasmid was also circulating in the broader community but was lower in prevalence than in the jail, suggesting that there may be differing or stronger selective pressures for strains with the plasmid in the jail. In each chapter, the integration of genomics and epidemiology was crucial to our understanding of MRSA adaptation and spread.

5.1.1 Contributions to our understanding of USA300 MRSA epidemiology

In Chapter 2, we present a genomic epidemiological study of all MRSA clinical cultures presenting to Cook County Health over a 4-year period with detailed healthcare exposures of all patients across the state of Illinois. This work provided a comprehensive profile of MRSA epidemiology in the urban community of Chicago in recent years and allowed us to pinpoint or rule out focal points of USA300 MRSA spread.

Infection control in the healthcare setting had been successful at reducing MRSA infection incidence in the past decade, but the lack of similar decline in the community should shift focus outside of the healthcare setting [64, 65]. Indeed, we observe a lack of healthcare overlap among isolates involved in potential recent transmission on the statewide facility level, even among so-called "hospital-onset" and "healthcare-associated" USA300 infections (Chapter 2). We question the ability of these definitions in predicting sites of MRSA acquisition and hypothesize that infections that occur after the 72-hour threshold may likely be due to pre-existing colonization from the community. We suggest that continued use of these epidemiological definitions should not detract our focus away from pinpointing important hubs of transmission in the community and designing novel community interventions.

Moreover, this dissertation demonstrates the utility of genomics to add clarity to epidemiological studies. For example, individuals had numerous exposures to the healthcare setting before and after their MRSA diagnosis, which alone could implicate the healthcare setting as a source of acquisition and further spread. However, identifying putative transmission by way of genomics allowed us to determine that there was little direct or indirect overlap among individuals in the healthcare setting, suggesting that they acquired MRSA in the community.

5.1.2 Jails as amplifiers of MRSA transmission in the community

In Chapter 3, we zoomed in on one section of the community with a higher burden of MRSA – jails. Jails and prisons were one of the first places USA300 MRSA outbreaks in the community were observed. Jails are characterized by short stays whereas prisons have a more stable, long-term population of inmates. Thus, it has been speculated that smaller populations and shorter stays in jails compared to prisons may result in less transmission opportunity in jails [78]. However, in Chapter 3, we showed that within 30 days there was an ~8% acquisition rate of MRSA colonization. Moreover, we used genomics to uncover the existence of transmission among jail-onset infections mediated by location sharing and environmental contamination. This rate of transmission is a function of the high burden of MRSA colonization at intake to the jail: 19% compared to the national average of 1.5% [7].

As jails have a high rate of turnover back into the community, interventions in the jails may help reduce communicable diseases to the community at large. This has been demonstrated in syphilis, where jail-based interventions for syphilis have resulted in reduced rates of syphilis in the surrounding community [111]. Efforts to control MRSA spread in jails have been successful and include a combination of skin lesion screening, standardized antimicrobial therapy, wound care, and enhanced hygiene including hand hygiene and chlorhexidine body wash [112]. Our results suggest that enhanced environmental cleaning particularly in locations where individuals who have long stays have resided, further investigation of certain buildings, and enhanced hygiene or decolonization of those at entry could be effective interventions for MRSA. Further work is needed to assess the impact of these interventions on the downstream community burden of MRSA, paying attention to both the burden of MRSA colonization and infection. One way to assess this specifically could be to monitor changes in the community of strains that are particularly prevalent in the jail, such as the presence of the plasmid we observed in Chapter 4.

The jail also may be a good place to surveil for strains that may emerge to cause success

in the community. In Chapter 4, we show evidence of multiple acquisitions of a plasmid and subsequent spread in the jail, with ~80% of infected individuals involved in transmission pairs carrying the plasmid (compared to ~30% not involved in transmission). We see this plasmid in the community, increasing in prevalence in time, and associated with incarceration; further work is needed to assess the role of the jail in the introduction of new strains in the broader community. Another study has shown that prisons select for multidrug resistant variants in *Mycobacterium tuberculosis* that lack the typical fitness costs associated with increased resistance, and these strains overflow into the community [55]. Thus it is important to monitor emerging variants in correctional facilities and prevent their further spread in the community.

5.1.3 A framework for genomic epidemiology in jails

Genomic epidemiology is a relatively new field that was pioneered in the healthcare setting to retrospectively detect outbreaks [41]. There are far fewer genomic epidemiological studies in the community and even fewer in jails/prisons in part because of challenges of data collection. Our collaborators were able to overcome these challenges to collect bacterial samples and detailed metadata in the Cook County Jail. Insights into the data collection process that we gleaned from the post-collection analysis are provided below.

Considerations for collection and sequencing of microbial isolates

Considerations for sample collection include culture type (e.g. surveillance versus clinical cultures), body site of collection, and number of samples to collect. As MRSA and other bacterial pathogens can both colonize and infect, and both states play a role in transmission, collecting both colonization and clinical infections will maximize transmission linkage detection. However, we found that if transmission is common enough, many transmission links can be detected with just clinical cultures.

Furthermore, in Chapter 3, we collected colonization samples from the nose, throat, and groin. As MRSA colonization is often extra-nasal, without collecting samples from multiple body sites we may have missed numerous transmission links. A meta-analysis of studies conducting extra-nasal and nasal screening for MRSA at hospital or ICU encounters found that extra-nasal screening identified one-third more MRSA cases than nasal screening alone [113]. Moreover, collection of multiple isolates per person can serve as a quality control for the sequencing process, as we found that most individuals are colonized with the same strain across body site [67]. Although most strains across body sites are concordant, collecting from multiple body sites can contribute to capturing the diversity of MRSA colonizing the host [62]. Another consideration for capturing the diversity of MRSA in the host the number of colonies to sequence per body site. In all studies presented in this dissertation, we sequenced one bacterial colony per body site. Collecting one colony per site will limit capturing the full diversity present on an individual and may effect reconstruction of transmission networks [114]. However, collecting from multiple body sites mitigates this concern. Further, we were able to detect numerous transmission events with epidemiological support using the single colony approach [67]. With limited funding, there is a trade-off between density of sampling of a single individual and the number of people in your study. For the purposes of constructing transmission networks, we prioritized the latter.

Considerations for metadata collection

We collected metadata about factors related to behaviors in the broader community (e.g. drug use, zipcode) and in the jail itself (e.g. sharing personal items, using the gym). Because of this, we were able to identify both community transmission networks [38] and factors associated with transmission within the jail (Chapter 3). Collection of granular location data is also important. In the healthcare setting, multiple levels of location can be collected such as facility, floor, ward, and room [115]. Here, we also collected multiple levels of the

jail which included building and living unit (e.g. cell or dorm room). The most granular level data (i.e. living unit) gave us confidence that we detected direct transmission among closely related genetic pairs and that there may be a role of environmental contamination if two individuals with closely related MRSA lived in the same room separated by time. The least granular data (i.e. building) allowed us to identify where MRSA transmission was most common and that both cell and dorm-based buildings had high levels of transmission.

While preparing this dissertation, SARS-CoV-2 emerged to cause a global pandemic. Jails, including the Cook County Jail, were particularly affected [92]. The sample collection infrastructure, collaborative relationships, and analytic framework described in this dissertation afforded a quick turnover time to assess SARS-CoV-2 transmission in the Cook County and broader community (unpublished data).

5.1.4 The importance of antimicrobial stewardship

In Chapter 4, we identified an association between involvement in a transmission pair and harboring MRSA with a plasmid that carries the gene *ermC* which confers resistance to the macrolide and lincosamide classes of antibiotics. This association was stronger in the jail setting, but also present in a comprehensive sampling of clinical cultures. We speculate that the association of the plasmid with transmission is a result of either or a combination of 1) an increased fitness under antimicrobial pressures or 2) fitness advantages gained from a change in the cellular proteome as a result of methylation of the ribosomal RNA by ErmC [116]. The latter could be tested *in vitro* with RNA sequencing experiments comparing strains with and without the plasmid under differing conditions including rich and minimal media and under antibiotic stress.

Regarding the first hypothesis, in the clinical setting, macrolide and lincosamide exposure in the past 6 months were not significantly associated with harboring the *ermC*-carrying plasmid; however, over 50% of individuals in the study had taken the macrolide clindamycin

in the past 6 months suggesting a plausible role for clindamycin use in the success of the strains harboring the plasmid. Another antimicrobial pressure could be from the additive triclosan, a chemical added to many consumer products including soaps, deodorant, plastics, and textiles for its antibacterial properties [100, 101]. A study of athletic and educational facilities revealed a correlation between *erm* genes and the concentration of triclosan [102]. Further, triclosan has been shown to promote nasal colonization of *S. aureus*. Triclosan is common in many products and as a result, one study of the general U.S. population found triclosan in about 75% of people in the study [100]. We wonder if the increased prevalence of strains harboring the plasmid could be a result of differing levels of triclosan exposure such as hygiene or cleaning products provided in the jail.

The potential for selection of a transmissible variant by antibiotic or antmicrobial products highlights the importance of antimicrobial stewardship in healthcare settings and beyond. Antimicrobial stewardship efforts in hospital setting have largely been focused on antibiotics [117]. Reducing transmission through environmental cleaning is one strategy to reduce the need for antibiotics to treat infections. However, these results highlight the potential to consider the regulation of products containing antimicrobial additives such as triclosan. Further, this work highlights the importance of genomic surveillance of strains in real-time to inform if changes in antimicrobial practices are warranted.

5.2 The future of genomic epidemiology

This dissertation demonstrates how genomics can be used retrospectively to assess transmission and variant emergence. The future beholds use of real-time use of genomics in a clinical setting, in remote locations, and hopefully in community settings such as jails. The analytical framework presented lays a piece of the foundation for this future. Reviewed below are the new methodologies, technologies, and insights from related research that will move us

toward routine use of genomics in clinical and infection control decision-making.

5.2.1 Moving away from SNV thresholds?

One challenge in genomic epidemiology is the accurate identification of transmission pairs. Currently, a common approach to infer transmission is to use a single SNV threshold. SNV thresholds vary by organism and depend on the evolutionary rate, though no pan-organism algorithm exists to calculate SNV thresholds. Use of a single SNV threshold can result in identification of false positive or false negative transmission links. In MRSA, a SNV threshold of 40 has been historically used as it was the maximum within-diversity observed in an individual and therefore the full diversity that could be transmitted. Interestingly, studies have taken different approaches to determine the optimum SNV threshold for MRSA and independently settled on a similar number [71, 70]. We addressed the challenges of a single SNV threshold by conducting a sensitivity analysis over multiple SNV thresholds (Chapter 3).

Recent efforts have proposed phylogenetic-based approaches as alternatives to SNV thresholds. So far, these approaches require either comprehensive sampling of a facility [115] or collecting multiple samples per individual [71]. Both approaches can be difficult logistically in the study design and expensive to put into practice. However, one study showed that when MRSA was routinely surveyed, multiple samples were not needed to detect transmission in an outbreak setting [118]. More work is needed to critically assess how much more information a SNV threshold-independent approach can add to epidemiological studies while considering the feasibility in a real-time setting.

5.2.2 Real-time outbreak detection

Genomic epidemiology has been used retroactively to investigate many outbreaks in the healthcare setting [42]. The ever decreasing price of whole-genome sequencing could make surveillance and real-time detection of outbreaks a possibility. However, obstacles for routine use of real-time use of sequencing technology include a need for expensive equipment and training of personnel for library preparation and processing, interpreting, and storage of sequencing data [119, 120]. Incremental progress has been made on both fronts through the development of portable, affordable sequencing technology and development of bioinformatics protocols and proof-of-concept studies for use in a clinical setting.

Technological advances like the MinION from Oxford Nanopore Technologies has provided portable DNA sequencing devices that can be easily transported to the field with rapid sequencing and downstream analysis [119, 121]. This technology has been deployed in remote settings in response to the Ebola [122] and Zika [123] outbreaks. Still, limitations exist in cost of flow cell, the need for PCR amplification before sequencing, and error rates of variant calls if not sequenced deeply enough [121].

From a logistical standpoint, at present, trained bioinformatics personnel will be needed for implementation of genomics in a clinical lab. Studies have aimed to develop protocols and automate sequencing processing pipelines. For example, in a clinical setting, a protocol was developed and benchmarked against a previously investigated outbreak of MRSA using Oxford Nanopore Technologies and subsequently used to identify two S. aureus outbreaks in less than 31 hours [119]. Another study implemented a fully-automated bioinformatics tool in a clinical microbiology lab to predict antibiotic resistance from the genome [124]. Still, more work to automate and standardize bioinformatics pipelines across clinical labs and to train personnel will be needed to make real-time genomics a reality on a global scale.

This dissertation demonstrates that MRSA surveillance and interventions to reduce MRSA transmission in real-time would be useful, and genomics would be required over

molecular type methods to detect an outbreak as most strains entering the jail were USA300. Proof-of-concept studies and protocols should be developed for community settings, such as jails.

5.2.3 Precision treatment of infection from bacterial genomes

As the price of whole-genome sequencing declines and more becomes known about the genetic determinants of resistance and virulence, sequencing has the potential to move into routine clinical practice as an alternative to phenotypic antibiotic susceptibility testing and to provide precision treatment to patients. Already, genomics has been used clinically to determine resistance profiles in *Mycobacterium tuberculosis*, and it has been shown to be a more affordable and time-efficient alternative to susceptibility testing in this slow-growing bacteria [125, 126]. Much work has been done in MRSA to predict genetic determinants of resistance from the genome [127, 128]. However, even with databases of known resistance elements, another barrier is the informatics expertise in clinical labs to process sequencing data. Here too though, automated bioinformatics pipelines to test antibiotic resistance in MRSA have been tested in the clinical setting with both logistical success and concordance with susceptibility testing [124].

A more forward-thinking application of whole-genome sequencing is the use of genomics in optimizing treatment strategies. Work has been done in MRSA to try to predict virulence (e.g. toxicity) from the genome [129]. In theory, patients with more virulent lineages could be quarantined to prevent transmission, monitored more closely for complications, or treated with anti-virulence drugs in a future where those exist [129].

5.3 Conclusion

While MRSA has declined in the past decade, further reduction in cases will require community-based interventions. Indeed, we identified a lack of healthcare transmission among USA300 MRSA (Chapter 2). We identified potential points of intervention to reduce transmission in the jail (Chapter 3) and identified the existence of potential selective pressure enhancing transmission in the jail (Chapter 4). Further work will be needed to determine the extent that the jail amplifies community MRSA transmission and how jail-based interventions impact the burden of MRSA in the community. Moreover, additional community-based genomic epidemiological studies are needed to identify other hubs of MRSA transmission. Finally, this work provided an analytical framework for future genomic epidemiology studies, and has already been useful in studying the COVID-19 pandemic in the backdrop of the Cook County Jail.

Bibliography

[1] Chris Woolston. "Graduate survey: A love–hurt relationship". In: *Nature* 550.7677 (Oct. 2017). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7677 Primary_atype: Special Features Publisher: Nature Publishing Group Subject_term: Careers Subject_term_id: careers, pp. 549–552. ISSN: 1476-4687. DOI: 10.1038/nj7677-549a. URL: http://www.nature.com/articles/nj7677-549a.

[2] John N. Sheagren. "Staphylococcus aureus". In: *New England Journal of Medicine* 310.21 (May 24, 1984). Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/N pp. 1368–1373. ISSN: 0028-4793. DOI: 10.1056/NEJM198405243102107. URL: https://doi.org/10.1056/NEJM198405243102107.

[3] Henry F. Chambers and Frank R. Deleo. "Waves of resistance: Staphylococcus aureus in the antibiotic era". In: *Nat Rev Microbiol* 7.9 (Sept. 2009), pp. 629–641. ISSN: 1740-1534. DOI: 10.1038/nrmicro2200.

[4] Catriona P. Harkins et al. "Methicillin-resistant Staphylococcus aureus emerged long before the introduction of methicillin into clinical practice". In: *Genome Biol* 18 (July 20, 2017). ISSN: 1474-7596. DOI: 10.1186/s13059-017-1252-9. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5517843/.

[5] International Working Group on the Classification of Staphylococcal Cassette Chromosome and Elements (IWG-SCC)*. "Classification of Staphylococcal Cassette Chromosome *mec* (SCC *mec*): Guidelines for Reporting Novel SCC *mec* Elements". In:

Antimicrob Agents Chemother 53.12 (Dec. 2009), pp. 4961–4967. ISSN: 0066-4804, 1098-6596. DOI: 10.1128/AAC.00579-09. URL: https://journals.asm.org/doi/10.1128/AAC.00579-09.

[6] Ari Robicsek, Jennifer L. Beaumont, and Lance R. Peterson. "Duration of Colonization with Methicillin-Resistant Staphylococcus aureus". In: *Clinical Infectious Diseases* 48.7 (Apr. 1, 2009), pp. 910–913. ISSN: 1058-4838. DOI: 10.1086/597296. URL: https://doi.org/10.1086/597296.

[7] Rachel J. Gorwitz et al. "Changes in the prevalence of nasal colonization with Staphylococcus aureus in the United States, 2001-2004". In: *J Infect Dis* 197.9 (May 1, 2008), pp. 1226–1234. ISSN: 0022-1899. DOI: 10.1086/533494.

[8] Nancy F. Crum-Cianflone et al. "Association of Methicillin-Resistant Staphylococcus aureus (MRSA) Colonization With High-Risk Sexual Behaviors in Persons Infected With Human Immunodeficiency Virus (HIV)". In: *Medicine* 90.6 (Nov. 2011), pp. 379–389. ISSN: 0025-7974. DOI: 10.1097/MD.0b013e318238dc2c. URL: https://journals.lww.com/md-journal/Fulltext/2011/11000/Association_of_Methicillin_Resistant.3.aspx.

[9] Selda Sayin Kutlu et al. "Prevalence and risk factors for methicillin-resistant Staphylococcus aureus colonization in a diabetic outpatient population: A prospective cohort study". In: *American Journal of Infection Control* 40.4 (May 1, 2012). Publisher: Elsevier, pp. 365–368. ISSN: 0196-6553, 1527-3296. DOI: 10.1016/j.ajic.2011.05.009. URL: https://www.ajicjournal.org/article/S0196-6553(11)00831-5/abstract.

[10] Jessica H. Leibler et al. "Prevalence and risk factors for MRSA nasal colonization among persons experiencing homelessness in Boston, MA". In: *Journal of Medical Microbiology* 66.8 (Aug. 1, 2017), pp. 1183–1188. ISSN: 0022-2615, 1473-5644. DOI:

10.1099/jmm.0.000552. URL: https://www.microbiologyresearch.org/content/journal/jmm/10.1099/jmm.0.000552.

[11] Kyle J. Popovich et al. "Community-Associated Methicillin-Resistant Staphylococcus aureus Colonization Burden in HIV-Infected Patients". In: *Clinical Infectious Diseases* 56.8 (Apr. 15, 2013), pp. 1067–1074. ISSN: 1058-4838. DOI: 10.1093/cid/cit010. URL: https://doi.org/10.1093/cid/cit010.

[12] C. von Eiff et al. "Nasal carriage as a source of Staphylococcus aureus bacteremia. Study Group". In: *N Engl J Med* 344.1 (Jan. 4, 2001), pp. 11–16. ISSN: 0028-4793. DOI: 10.1056/NEJM200101043440102.

[13] Centers for Disease Control and Prevention (U.S.) *Antibiotic resistance threats in the United States, 2019*. Centers for Disease Control and Prevention (U.S.), Nov. 2019. DOI: 10.15620/cdc:82532. URL: https://stacks.cdc.gov/view/cdc/82532.

[14] Hajo Grundmann et al. "Emergence and resurgence of meticillin-resistant Staphylococcus aureus as a public-health threat". In: *The Lancet* 368.9538 (Sept. 2, 2006), pp. 874–885. ISSN: 0140-6736. DOI: 10.1016/S0140-6736(06)68853-3. URL: https://www.sciencedirect.com/science/article/pii/S0140673606688533.

[15] Sara E. Cosgrove et al. "Comparison of Mortality Associated with Methicillin-Resistant and Methicillin-Susceptible Staphylococcus aureus Bacteremia: A Meta-analysis". In: *Clinical Infectious Diseases* 36.1 (Jan. 1, 2003), pp. 53–59. ISSN: 1058-4838. DOI: 10.1086/345476. URL: https://doi.org/10.1086/345476.

[16] F. Rozgonyi et al. "Is MRSA more virulent than MSSA?" In: *Clinical Microbiology and Infection* 13.9 (2007). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-0691.2007.01780.x, pp. 843–845. ISSN: 1469-0691. DOI: https://doi.org/10.1111/j.1469-0691.2007.01780.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-0691.2007.01780.x.

[17] Loren G. Miller et al. "A prospective investigation of outcomes after hospital discharge for endemic, community-acquired methicillin-resistant and -susceptible Staphylococcus aureus skin infection". In: *Clin Infect Dis* 44.4 (Feb. 15, 2007), pp. 483–492. ISSN: 1537-6591. DOI: 10.1086/511041.

[18] Rachel J. Gordon and Franklin D. Lowy. "Pathogenesis of Methicillin-Resistant Staphylococcus aureus Infection". In: *Clinical Infectious Diseases* 46 (Supplement_5 June 1, 2008), S350–S359. ISSN: 1058-4838. DOI: 10.1086/533591. URL: https://doi.org/10.1086/533591.

[19] Min Li et al. "Evolution of virulence in epidemic community-associated methicillin-resistant Staphylococcus aureus". In: *PNAS* 106.14 (Apr. 7, 2009), pp. 5883–5888. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0900743106. URL: http://www.pnas.org/content/106/14/5883.

[20] Eun Ju Choo and Henry F. Chambers. "Treatment of Methicillin-Resistant Staphylococcus aureus Bacteremia". In: *Infect Chemother* 48.4 (Dec. 2016), pp. 267–273. ISSN: 2093-2340. DOI: 10.3947/ic.2016.48.4.267. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5204005/.

[21] Will A. McGuinness, Natalia Malachowa, and Frank R. DeLeo. "Vancomycin Resistance in Staphylococcus aureus". In: *Yale J Biol Med* 90.2 (June 23, 2017), pp. 269–281. ISSN: 0044-0086. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5482303/.

[22] B. Shopsin et al. "Evaluation of protein A gene polymorphic region DNA sequencing for typing of Staphylococcus aureus strains". In: *J Clin Microbiol* 37.11 (Nov. 1999), pp. 3556–3563. ISSN: 0095-1137. DOI: 10.1128/JCM.37.11.3556-3563.1999.

[23] Linda K. McDougal et al. "Pulsed-Field Gel Electrophoresis Typing of Oxacillin-Resistant Staphylococcus aureus Isolates from the United States: Establishing a National Database". In: *J Clin Microbiol* 41.11 (Nov. 2003), pp. 5113–5120. ISSN: 0095-

1137. DOI: 10.1128/JCM.41.11.5113-5120.2003. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC262524/.

[24] Mark C. Enright et al. "Multilocus Sequence Typing for Characterization of Methicillin-Resistant and Methicillin-Susceptible Clones of Staphylococcus aureus". In: *J Clin Microbiol* 38.3 (Mar. 2000), pp. 1008–1015. ISSN: 0095-1137. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC86325/.

[25] Jolene R. Bowers et al. "Improved Subtyping of Staphylococcus aureus Clonal Complex 8 Strains Based on Whole-Genome Phylogenetic Analysis". In: *mSphere* 3.3 (June 27, 2018). Publisher: American Society for Microbiology Journals Section: Research Article. ISSN: 2379-5042. DOI: 10.1128/mSphere.00464-17. URL: https://msphere.asm.org/content/3/3/e00464-17.

[26] Centers for Disease Control and Prevention (CDC). "Methicillin-resistant Staphylococcus aureus skin or soft tissue infections in a state prison–Mississippi, 2000". In: *MMWR Morb Mortal Wkly Rep* 50.42 (Oct. 26, 2001), pp. 919–922. ISSN: 0149-2195.

[27] Centers for Disease Control and Prevention (CDC). "Methicillin-resistant Staphylococcus aureus infections in correctional facilities—Georgia, California, and Texas, 2001-2003". In: *MMWR Morb Mortal Wkly Rep* 52.41 (Oct. 17, 2003), pp. 992–996. ISSN: 1545-861X.

[28] Centers for Disease Control and Prevention (CDC). "Four pediatric deaths from community-acquired methicillin-resistant Staphylococcus aureus — Minnesota and North Dakota, 1997-1999". In: *MMWR Morb Mortal Wkly Rep* 48.32 (Aug. 20, 1999), pp. 707–710. ISSN: 0149-2195.

[29] Sophia V Kazakova et al. "A Clone of Methicillin-Resistant Staphylococcus aureus among Professional Football Players". In: *The New England Journal of Medicine* (2005), p. 8.

[30] Michael Z. David and Robert S. Daum. "Community-Associated Methicillin-Resistant Staphylococcus aureus: Epidemiology and Clinical Consequences of an Emerging Epidemic". In: *Clin Microbiol Rev* 23.3 (July 2010), pp. 616–687. ISSN: 0893-8512. DOI: 10.1128/CMR.00081-09. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2901661/.

[31] Fred C. Tenover and Richard V. Goering. "Methicillin-resistant Staphylococcus aureus strain USA300: origin and epidemiology". In: *Journal of Antimicrobial Chemotherapy* 64.3 (Sept. 1, 2009), pp. 441–446. ISSN: 0305-7453. DOI: 10.1093/jac/dkp241. URL: https://doi.org/10.1093/jac/dkp241.

[32] Gregory J. Moran et al. "Methicillin-resistant S. aureus infections among patients in the emergency department". In: *N Engl J Med* 355.7 (Aug. 17, 2006), pp. 666–674. ISSN: 1533-4406. DOI: 10.1056/NEJMoa055356.

[33] Kyle J. Popovich, Robert A. Weinstein, and Bala Hota. "Are Community-Associated Methicillin-Resistant Staphylococcus aureus (MRSA) Strains Replacing Traditional Nosocomial MRSA Strains?" In: *Clinical Infectious Diseases* 46.6 (Mar. 15, 2008), pp. 787–794. ISSN: 1058-4838. DOI: 10.1086/528716. URL: https://doi.org/10.1086/528716.

[34] Kyle J. Popovich et al. "Genomic and Epidemiological Evidence for Community Origins of Hospital-Onset Methicillin-Resistant Staphylococcus aureus Bloodstream Infections". In: *The Journal of Infectious Diseases* 215.11 (June 1, 2017), pp. 1640–1647. ISSN: 0022-1899. DOI: 10.1093/infdis/jiw647. URL: https://doi.org/10.1093/infdis/jiw647.

[35] Kyle J Popovich. "Intersection of HIV and community-associated methicillin-resistant Staphylococcus aureus". In: *Future Virology* 15.1 (Jan. 1, 2020). Publisher: Future

Medicine, pp. 53–65. ISSN: 1746-0794. DOI: 10.2217/fvl-2019-0093. URL: https://www.futuremedicine.com/doi/full/10.2217/fvl-2019-0093.

[36] Greg A. Greenberg and Robert A. Rosenheck. "Jail incarceration, homelessness, and mental health: a national study". In: *Psychiatr Serv* 59.2 (Feb. 2008), pp. 170–177. ISSN: 1075-2730. DOI: 10.1176/ps.2008.59.2.170.

[37] Ryan P. Westergaard, Anne C. Spaulding, and Timothy P. Flanigan. "HIV among persons incarcerated in the US: a review of evolving concepts in testing, treatment and linkage to community care". In: *Curr Opin Infect Dis* 26.1 (Feb. 2013), pp. 10–16. ISSN: 0951-7375. DOI: 10.1097/QCO.0b013e32835c1dd0. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3682655/.

[38] Kyle J. Popovich et al. "Frequent Methicillin-Resistant Staphylococcus aureus Introductions Into an Inner-city Jail: Indications of Community Transmission Networks". In: *Clin Infect Dis* 71.2 (July 11, 2020), pp. 323–331. ISSN: 1537-6591. DOI: 10.1093/cid/ciz818.

[39] Kyle J. Popovich et al. "Community-associated methicillin-resistant Staphylococcus aureus and HIV: intersecting epidemics". In: *Clin Infect Dis* 50.7 (Apr. 1, 2010), pp. 979–987. ISSN: 1537-6591. DOI: 10.1086/651076.

[40] Bala Hota et al. "Community-associated methicillin-resistant Staphylococcus aureus skin and soft tissue infections at a public hospital: do public housing and incarceration amplify transmission?" In: *Arch Intern Med* 167.10 (May 28, 2007), pp. 1026–1033. ISSN: 0003-9926. DOI: 10.1001/archinte.167.10.1026.

[41] Evan S. Snitkin et al. "Tracking a Hospital Outbreak of Carbapenem-Resistant Klebsiella pneumoniae with Whole-Genome Sequencing". In: *Sci Transl Med* 4.148 (Aug. 22, 2012), 148ra116. ISSN: 1946-6234. DOI: 10.1126/scitranslmed.3004129. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3521604/.

[42] Kyle J. Popovich and Evan S. Snitkin. "Whole Genome Sequencing—Implications for Infection Prevention and Outbreak Investigations". In: *Curr Infect Dis Rep* 19.4 (Mar. 9, 2017), p. 15. ISSN: 1534-3146. DOI: 10.1007/s11908-017-0570-0. URL: https://doi.org/10.1007/s11908-017-0570-0.

[43] Simon R Harris et al. "Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study". In: *Lancet Infect Dis* 13.2 (Feb. 2013), pp. 130–136. ISSN: 1473-3099. DOI: 10.1016/S1473-3099(12)70268-2. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3556525/.

[44] Anne-Catrin Uhlemann et al. "Molecular tracing of the emergence, diversification, and transmission of S. aureus sequence type 8 in a New York community". In: *Proc Natl Acad Sci U S A* 111.18 (May 6, 2014), pp. 6738–6743. ISSN: 0027-8424. DOI: 10.1073/pnas.1401006111. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4020051/.

[45] Emily Kajita et al. "Modelling an outbreak of an emerging pathogen". In: *Nat Rev Microbiol* 5.9 (Sept. 2007). Number: 9 Publisher: Nature Publishing Group, pp. 700–709. ISSN: 1740-1534. DOI: 10.1038/nrmicro1660. URL: https://www.nature.com/articles/nrmicro1660.

[46] Jacques Baillargeon et al. "Methicillin-Resistant Staphylococcus aureus Infection in the Texas Prison System". In: *Clinical Infectious Diseases* 38.9 (May 1, 2004), e92–e95. ISSN: 1058-4838. DOI: 10.1086/383146. URL: https://doi.org/10.1086/383146.

[47] Michael Z. David et al. "Predominance of Methicillin-Resistant Staphylococcus aureus among Pathogens Causing Skin and Soft Tissue Infections in a Large Urban Jail: Risk Factors and Recurrence Rates". In: *J Clin Microbiol* 46.10 (Oct. 2008), pp. 3222–3227. ISSN: 0095-1137. DOI: 10.1128/JCM.01423-08. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2566069/.

[48] Paul J. Planet et al. "Emergence of the Epidemic Methicillin-Resistant Staphylococcus aureus Strain USA300 Coincides with Horizontal Transfer of the Arginine Catabolic Mobile Element and speG-mediated Adaptations for Survival on Skin". In: *mBio* 4.6 (Dec. 17, 2013). ISSN: 2150-7511. DOI: 10.1128/mBio.00889-13. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3870260/.

[49] Lavanya Challagundla et al. "Range Expansion and the Origin of USA300 North American Epidemic Methicillin-Resistant Staphylococcus aureus". In: *mBio* 9.1 (Mar. 7, 2018), e02016–17. ISSN: , 2150-7511. DOI: 10.1128/mBio.02016-17. URL: http://mbio.asm.org/content/9/1/e02016-17.

[50] Paul J. Planet. "Life After USA300: The Rise and Fall of a Superbug". In: *The Journal of Infectious Diseases* 215 (suppl_1 Feb. 15, 2017), S71–S77. ISSN: 0022-1899. DOI: 10.1093/infdis/jiw444. URL: https://doi.org/10.1093/infdis/jiw444.

[51] Philippe Glaser et al. "Demography and Intercontinental Spread of the USA300 Community-Acquired Methicillin-Resistant Staphylococcus aureus Lineage". In: *mBio* 7.1 (Feb. 16, 2016). ISSN: 2150-7511. DOI: 10.1128/mBio.02183-15. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4752609/.

[52] Solomon A. Mekonnen et al. "Signatures of cytoplasmic proteins in the exoproteome distinguish community- and hospital-associated methicillin-resistant Staphylococcus aureus USA300 lineages". In: *Virulence* 8.6 (Aug. 18, 2017). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/21505594.2017.1325064, pp. 891–907. ISSN: 2150-5594. DOI: 10.1080/21505594.2017.1325064. URL: https://doi.org/10.1080/21505594.2017.1325064.

[53] Solomon A. Mekonnen et al. "Metabolic niche adaptation of community- and hospital-associated methicillin-resistant Staphylococcus aureus". In: *Journal of Proteomics*

(Oct. 12, 2018). ISSN: 1874-3919. DOI: 10.1016/j.jprot.2018.10.005. URL: http://www.sciencedirect.com/science/article/pii/S1874391918303725.

[54] Binh An Diep et al. "Emergence of multidrug-resistant, community-associated, methicillin-resistant Staphylococcus aureus clone USA300 in men who have sex with men". In: *Ann Intern Med* 148.4 (Feb. 19, 2008), pp. 249–257. ISSN: 1539-3704. DOI: 10.7326/0003-4819-148-4-200802190-00204.

[55] Sebastian M. Gygli et al. "Prisons as ecological drivers of fitness-compensated multidrug-resistant Mycobacterium tuberculosis". In: *Nat Med* (May 24, 2021). Publisher: Nature Publishing Group, pp. 1–7. ISSN: 1546-170X. DOI: 10.1038/s41591-021-01358-x. URL: http://www.nature.com/articles/s41591-021-01358-x.

[56] Scott K. Fridkin et al. "Methicillin-resistant Staphylococcus aureus disease in three communities". In: *N Engl J Med* 352.14 (Apr. 7, 2005), pp. 1436–1444. ISSN: 1533-4406. DOI: 10.1056/NEJMoa043252.

[57] Centers for Disease Control and Prevention (CDC). "Methicillin-resistant staphylococcus aureus infections among competitive sports participants–Colorado, Indiana, Pennsylvania, and Los Angeles County, 2000-2003". In: *MMWR Morb Mortal Wkly Rep* 52.33 (Aug. 22, 2003), pp. 793–795. ISSN: 1545-861X.

[58] Craig E. Zinderman et al. "Community-acquired Methicillin-resistant Staphylococcus aureus among Military Recruits". In: *Emerg Infect Dis* 10.5 (May 2004), pp. 941–944. ISSN: 1080-6040. DOI: 10.3201/eid1005.030604. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3323224/.

[59] Betsy C. Herold. "Community-Acquired Methicillin-Resistant Staphylococcus aureus in Children With No Identified Predisposing Risk". In: *JAMA* 279.8 (Feb. 25, 1998), p. 593. ISSN: 0098-7484. DOI: 10.1001/jama.279.8.593. URL: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.279.8.593.

[60] Kyle J. Popovich and Robert A. Weinstein. "The Graying of Methicillin-Resistant Staphylococcus aureus". In: *Infection Control and Hospital Epidemiology* 30.1 (2009). Publisher: [Cambridge University Press, The Society for Healthcare Epidemiology of America], pp. 9–12. ISSN: 0899-823X. DOI: 10.1086/592709. URL: http://www.jstor.org/stable/10.1086/592709.

[61] R. Monina Klevens et al. "Community-associated Methicillin-resistant Staphylococcus aureus and Healthcare Risk Factors". In: *Emerg Infect Dis* 12.12 (Dec. 2006), pp. 1991–1993. ISSN: 1080-6040. DOI: 10.3201/eid1212.060505. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3291355/.

[62] Tanya Golubchik et al. "Within-Host Evolution of Staphylococcus aureus during Asymptomatic Carriage". In: *PLOS ONE* 8.5 (May 1, 2013). Publisher: Public Library of Science, e61319. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0061319. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061319.

[63] Michael Z. David et al. "Determining Whether Methicillin-Resistant Staphylococcus aureus Is Associated With Health Care". In: *JAMA* 299.5 (Feb. 6, 2008), pp. 519–520. ISSN: 0098-7484. DOI: 10.1001/jama.299.5.519-a. URL: https://doi.org/10.1001/jama.299.5.519-a.

[64] Raymund Dantes et al. "National burden of invasive methicillin-resistant Staphylococcus aureus infections, United States, 2011". In: *JAMA Intern Med* 173.21 (Nov. 25, 2013), pp. 1970–1978. ISSN: 2168-6114. DOI: 10.1001/jamainternmed.2013.10423.

[65] Isaac See et al. "Trends in Incidence of Methicillin-resistant Staphylococcus aureus Bloodstream Infections Differ by Strain Type and Healthcare Exposure, United States, 2005-2013". In: *Clin Infect Dis* 70.1 (Jan. 1, 2020), pp. 19–25. ISSN: 1537-6591. DOI: 10.1093/cid/ciz158.

[66] Susan S. Huang et al. "Decolonization to Reduce Postdischarge Infection Risk among MRSA Carriers". In: *New England Journal of Medicine* 380.7 (Feb. 14, 2019). Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa1716771, pp. 638–650. ISSN: 0028-4793. DOI: 10.1056/NEJMoa1716771. URL: https://doi.org/10.1056/NEJMoa1716771.

[67] Kyle J Popovich et al. "Genomic Epidemiology of MRSA During Incarceration at a Large Inner-City Jail". In: *Clinical Infectious Diseases* (ciaa1937 Jan. 4, 2021). ISSN: 1058-4838. DOI: 10.1093/cid/ciaa1937. URL: https://doi.org/10.1093/cid/ciaa1937.

[68] Timothy S. Naimi et al. "Comparison of community- and health care-associated methicillin-resistant Staphylococcus aureus infection". In: *JAMA* 290.22 (Dec. 10, 2003), pp. 2976–2984. ISSN: 1538-3598. DOI: 10.1001/jama.290.22.2976.

[69] Martin Hunt et al. "ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads". In: *Microb Genom* 3.10 (Sept. 4, 2017). ISSN: 2057-5858. DOI: 10.1099/mgen.0.000131. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5695208/.

[70] Marko Järvenpää et al. "A Bayesian model of acquisition and clearance of bacterial colonization incorporating within-host variation". In: *PLOS Computational Biology* 15.4 (Apr. 22, 2019). Publisher: Public Library of Science, e1006534. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1006534. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006534.

[71] Matthew D Hall et al. "Improved characterisation of MRSA transmission using within-host bacterial sequence diversity". In: *eLife* 8 (Oct. 8, 2019). Ed. by Mark Jit et al. Publisher: eLife Sciences Publications, Ltd, e46402. ISSN: 2050-084X. DOI: 10.7554/eLife.46402. URL: https://doi.org/10.7554/eLife.46402.

[72] Subha Kalyaanamoorthy et al. “ModelFinder: fast model selection for accurate phylogenetic estimates”. In: *Nat. Methods* 14.6 (June 2017), pp. 587–589. ISSN: 1548-7105. DOI: 10.1038/nmeth.4285.

[73] Lam-Tung Nguyen et al. “IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies”. In: *Mol. Biol. Evol.* 32.1 (Jan. 2015), pp. 268–274. ISSN: 1537-1719. DOI: 10.1093/molbev/msu300.

[74] Emmanuel Paradis and Klaus Schliep. “ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R”. In: *Bioinformatics* 35.3 (Feb. 1, 2019). Publisher: Oxford Academic, pp. 526–528. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bty633. URL: http://academic.oup.com/bioinformatics/article/35/3/526/5055127.

[75] Guangchuang Yu et al. “ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data”. In: *Methods in Ecology and Evolution* 8.1 (2017). _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12628, pp. 28–36. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12628. URL: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12628.

[76] Loren G. Miller and Binh An Diep. “Clinical practice: colonization, fomites, and virulence: rethinking the pathogenesis of community-associated methicillin-resistant Staphylococcus aureus infection”. In: *Clin Infect Dis* 46.5 (Mar. 1, 2008), pp. 752–760. ISSN: 1537-6591. DOI: 10.1086/526773.

[77] Kyle J. Popovich, Bala Hota, and Robert A. Weinstein. “Treatment of community-associated methicillin-resistant Staphylococcus aureus”. In: *Curr Infect Dis Rep* 9.5 (Sept. 2007), pp. 398–407. ISSN: 1523-3847. DOI: 10.1007/s11908-007-0062-8.

[78] Bianca Malcolm. “The Rise of Methicillin-Resistant Staphylococcus aureus in U.S. Correctional Populations”. In: *J Correct Health Care* 17.3 (July 2011), pp. 254–265.

ISSN: 1078-3458. DOI: 10.1177/1078345811401363. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3116074/.

[79] Centers for Disease Control and Prevention (CDC). "Outbreaks of community-associated methicillin-resistant Staphylococcus aureus skin infections–Los Angeles County, California, 2002-2003". In: *MMWR Morb Mortal Wkly Rep* 52.5 (Feb. 7, 2003), p. 88. ISSN: 0149-2195.

[80] Justin T. Okano and Sally Blower. "Are correctional facilities amplifying the epidemic of community-acquired methicillin-resistant Staphylococcus aureus?" In: *Nat Rev Microbiol* 8.1 (Jan. 2010), p. 83. ISSN: 1740-1534. DOI: 10.1038/nrmicro2200-c1.

[81] Kyle J Popovich et al. "1229. Prevalence and Acquisition of MRSA in Females During Incarceration at a Large Inner-City Jail". In: *Open Forum Infect Dis* 5 (Suppl 1 Nov. 26, 2018), S373. ISSN: 2328-8957. DOI: 10.1093/ofid/ofy210.1062. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6254776/.

[82] Jason E. Farley et al. "Prevalence, risk factors, and molecular epidemiology of methicillin-resistant Staphylococcus aureus among newly arrested men in Baltimore, Maryland". In: *Am J Infect Control* 36.9 (Nov. 2008), pp. 644–650. ISSN: 1527-3296. DOI: 10.1016/j.ajic.2008.05.005.

[83] Kyle J. Popovich et al. "Genomic Epidemiology of USA300 Methicillin-Resistant Staphylococcus aureus in an Urban Community". In: *Clin Infect Dis* 62.1 (Jan. 1, 2016), pp. 37–44. ISSN: 1537-6591. DOI: 10.1093/cid/civ794.

[84] Edwin D. Charlebois et al. "Population-based community prevalence of methicillin-resistant Staphylococcus aureus in the urban poor of San Francisco". In: *Clin Infect Dis* 34.4 (Feb. 15, 2002), pp. 425–433. ISSN: 1537-6591. DOI: 10.1086/338069.

[85] S. Blank et al. "New approaches to syphilis control. Finding opportunities for syphilis treatment and congenital syphilis prevention in a women's correctional setting". In: *Sex Transm Dis* 24.4 (Apr. 1997), pp. 218–226. ISSN: 0148-5717. DOI: 10.1097/00007435-199704000-00006.

[86] Centers for Disease Control and Prevention (CDC). "Syphilis screening among women arrestees at the Cook County Jail–Chicago, 1996". In: *MMWR Morb Mortal Wkly Rep* 47.21 (June 5, 1998), pp. 432–433. ISSN: 0149-2195.

[87] Pennan M. Barry et al. "Is jail screening associated with a decrease in Chlamydia positivity among females seeking health services at community clinics?-San francisco, 1997-2004". In: *Sex Transm Dis* 36.2 (Feb. 2009), S22–28. ISSN: 1537-4521. DOI: 10.1097/OLQ.0b013e31815ed7c8.

[88] Jennifer Broad et al. "The impact of discontinuation of male STD screening services at a large urban county jail: Chicago, 2002-2004". In: *Sex Transm Dis* 36.2 (Feb. 2009), S49–52. ISSN: 1537-4521. DOI: 10.1097/OLQ.0b013e318156159a.

[89] Cynthia L. Maree et al. "Risk factors for infection and colonization with community-associated methicillin-resistant Staphylococcus aureus in the Los Angeles County jail: a case-control study". In: *Clin Infect Dis* 51.11 (Dec. 1, 2010), pp. 1248–1257. ISSN: 1537-6591. DOI: 10.1086/657067.

[90] Kyle J. Popovich et al. "Community-associated methicillin-resistant Staphylococcus aureus colonization in high-risk groups of HIV-infected patients". In: *Clin Infect Dis* 54.9 (May 2012), pp. 1296–1303. ISSN: 1537-6591. DOI: 10.1093/cid/cis030.

[91] Nasia Safdar et al. "Comparison of culture screening methods for detection of nasal carriage of methicillin-resistant Staphylococcus aureus: a prospective study comparing 32 methods". In: *J Clin Microbiol* 41.7 (July 2003), pp. 3163–3166. ISSN: 0095-1137. DOI: 10.1128/jcm.41.7.3163-3166.2003.

[92] Chad Zawitz et al. "Outbreak of COVID-19 and Interventions in One of the Largest Jails in the United States — Cook County, IL, 2020". In: *medRxiv* (July 14, 2020). Publisher: Cold Spring Harbor Laboratory Press, p. 2020.07.12.20148494. DOI: 10.1101/2020.07.12.20148494. URL: https://www.medrxiv.org/content/10.1101/2020.07.12.20148494v1.

[93] Natalia Malachowa and Frank R. DeLeo. "Mobile genetic elements of Staphylococcus aureus". In: *Cell. Mol. Life Sci.* 67.18 (Sept. 1, 2010), pp. 3057–3071. ISSN: 1420-9071. DOI: 10.1007/s00018-010-0389-4. URL: https://doi.org/10.1007/s00018-010-0389-4.

[94] Artur J. Sabat et al. "Complete-genome sequencing elucidates outbreak dynamics of CA-MRSA USA300 (ST8- spa t008) in an academic hospital of Paramaribo, Republic of Suriname". In: *Sci Rep* 7.1 (Jan. 20, 2017). Number: 1 Publisher: Nature Publishing Group, p. 41050. ISSN: 2045-2322. DOI: 10.1038/srep41050. URL: https://www.nature.com/articles/srep41050.

[95] Fred C Tenover et al. "Characterization of a Strain of Community-Associated Methicillin-Resistant Staphylococcus aureus Widely Disseminated in the United States". In: *J. CLIN. MICROBIOL.* 44 (2006), p. 11.

[96] Jose Cadena et al. "Clindamycin-resistant methicillin-resistant Staphylococcus aureus: epidemiologic and molecular characteristics and associated clinical factors". In: *Diagnostic Microbiology and Infectious Disease* 74.1 (Sept. 1, 2012), pp. 16–21. ISSN: 0732-8893. DOI: 10.1016/j.diagmicrobio.2012.05.010. URL: https://www.sciencedirect.com/science/article/pii/S0732889312001915.

[97] David A. Talan et al. "Comparison of Staphylococcus aureus From Skin and Soft-Tissue Infections in US Emergency Department Patients, 2004 and 2008". In: *Clin Infect Dis* 53.2 (July 15, 2011), pp. 144–149. ISSN: 1058-4838. DOI: 10.1093/cid/cir308.

URL: https://academic.oup.com/cid/article/53/2/144/285752/Comparison-of-Staphylococcus-aureus-From-Skin-and.

[98] Isabella A. Tickler et al. "Continued expansion of USA300-like methicillin-resistant Staphylococcus aureus (MRSA) among hospitalized patients in the United States". In: *Diagnostic Microbiology and Infectious Disease* 88.4 (Aug. 1, 2017), pp. 342–347. ISSN: 0732-8893. DOI: 10.1016/j.diagmicrobio.2017.04.016. URL: https://www.sciencedirect.com/science/article/pii/S0732889317301451.

[99] Sarah Sansom, Emily Benedict, and Stephanie Thiede, Bala Hota, Alla Aroutcheva, Darjai Payne, Chad Zawitz, Evan Snitkin, Stefan Green, Robert Weinstein, and Kyle Popovich. "Genomic Update of Phenotypic Prediction Rule for Methicillin-Resistant Staphylococcus aureus (MRSA) USA300 Discloses Jail Transmission Networks with Increased Resistance". In: Accepted (2021).

[100] Antonia M. Calafat et al. "Urinary Concentrations of Triclosan in the U.S. Population: 2003–2004". In: *Environ Health Perspect* 116.3 (Mar. 2008), pp. 303–307. ISSN: 0091-6765. DOI: 10.1289/ehp.10768. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2265044/.

[101] Lisa M. Weatherly and Julie A. Gosse. "Triclosan Exposure, Transformation, and Human Health Effects". In: *J Toxicol Environ Health B Crit Rev* 20.8 (2017), pp. 447–469. ISSN: 1093-7404. DOI: 10.1080/10937404.2017.1399306. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6126357/.

[102] Erica M. Hartmann et al. "Antimicrobial Chemicals Are Associated with Elevated Antibiotic Resistance Genes in the Indoor Dust Microbiome". In: *Environ Sci Technol* 50.18 (Sept. 20, 2016), pp. 9807–9815. ISSN: 0013-936X. DOI: 10.1021/acs.est.6b00262. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5032049/.

[103] Daniel E. Carey and Patrick J. McNamara. "The impact of triclosan on the spread of antibiotic resistance in the environment". In: *Front Microbiol* 5 (Jan. 15, 2015). ISSN: 1664-302X. DOI: 10.3389/fmicb.2014.00780. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4295542/.

[104] Corey Westfall et al. "The Widely Used Antimicrobial Triclosan Induces High Levels of Antibiotic Tolerance In Vitro and Reduces Antibiotic Efficacy up to 100-Fold In Vivo". In: *Antimicrob Agents Chemother* 63.5 (May 2019). ISSN: 1098-6596. DOI: 10.1128/AAC.02312-18.

[105] Adnan K. Syed et al. "Triclosan Promotes Staphylococcus aureus Nasal Colonization". In: *mBio* 5.2 (Apr. 8, 2014). ISSN: 2150-7511. DOI: 10.1128/mBio.01015-13. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3993860/.

[106] Gerry Tonkin-Hill et al. "Producing polished prokaryotic pangenomes with the Panaroo pipeline". In: *Genome Biology* 21.1 (July 22, 2020), p. 180. ISSN: 1474-760X. DOI: 10.1186/s13059-020-02090-4. URL: https://doi.org/10.1186/s13059-020-02090-4.

[107] Jaime Huerta-Cepas et al. "eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences". In: *Nucleic Acids Res* 44 (Database issue Jan. 4, 2016), pp. D286–D293. ISSN: 0305-1048. DOI: 10.1093/nar/gkv1248. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4702882/.

[108] B. Shrestha and S. S. Rana. "D test: a simple test with big implication for Staphylococcus aureus macrolide-lincosamide-streptograminB resistance pattern". In: *Nepal Med Coll J* 16.1 (Sept. 2014), pp. 88–94. ISSN: 2676-1319.

[109] John A Lees et al. "pyseer: a comprehensive tool for microbial pangenome-wide association studies". In: *Bioinformatics* 34.24 (Dec. 15, 2018), pp. 4310–4312. ISSN:

1367-4803. DOI: 10.1093/bioinformatics/bty539. URL: https://doi.org/10.1093/bioinformatics/bty539.

[110] S. F. Altschul et al. "Basic local alignment search tool". In: *J Mol Biol* 215.3 (Oct. 5, 1990), pp. 403–410. ISSN: 0022-2836. DOI: 10.1016/S0022-2836(05)80360-2.

[111] Richard H. Kahn et al. "Screening for Syphilis in Arrestees: Usefulness for Community-wide Syphilis Surveillance and Control". In: *Sexually Transmitted Diseases* 29.3 (Mar. 2002), pp. 150–156. ISSN: 0148-5717. URL: https://journals.lww.com/stdjournal/Fulltext/2002/03000/Screening_for_Syphilis_in_Arrestees__Usefulness.5.aspx.

[112] Susan H. Wootton et al. "Intervention to Reduce the Incidence of Methicillin-Resistant Staphylococcus aureus Skin Infections in a Correctional Facility in Georgia". In: *Infection Control and Hospital Epidemiology* 25.5 (2004). Publisher: [Cambridge University Press, The Society for Healthcare Epidemiology of America], pp. 402–407. ISSN: 0899-823X. DOI: 10.1086/502413. URL: http://www.jstor.org/stable/10.1086/502413.

[113] James A. McKinnell et al. "Quantifying the Impact of Extranasal Testing of Body Sites for Methicillin-Resistant Staphylococcus aureus Colonization at the Time of Hospital or Intensive Care Unit Admission". In: *Infection Control & Hospital Epidemiology* 34.2 (Feb. 2013). Publisher: Cambridge University Press, pp. 161–170. ISSN: 0899-823X, 1559-6834. DOI: 10.1086/669095. URL: https://www.cambridge.org/core/journals/infection-control-and-hospital-epidemiology/article/abs/quantifying-the-impact-of-extranasal-testing-of-body-sites-for-methicillinresistant-staphylococcus-aureus-colonization-at-the-time-of-hospital-or-intensive-care-unit-admission/F276D9353A5455803A360175F2CA283E.

[114] Gavin K. Paterson et al. "Capturing the cloud of diversity reveals complexity and heterogeneity of MRSA carriage, infection and transmission". In: *Nat Commun* 6.1 (Mar. 27, 2015). Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research

Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Bacterial pathogenesis;Genetics research;Population genetics Subject_term_id: bacterial-pathogenesis;genetics-research;population-genetics, p. 6560. ISSN: 2041-1723. DOI: 10.1038/ncomms7560. URL: http://www.nature.com/articles/ncomms7560.

[115] Shawn E. Hawken et al. "A novel threshold-independent approach to genomic cluster analysis discloses persistent routes of KPC+ Klebsiella pneumoniae transmission in a long-term acute care hospital". In: *medRxiv* (Sept. 28, 2020). Publisher: Cold Spring Harbor Laboratory Press, p. 2020.09.26.20200097. DOI: 10.1101/2020.09.26.20200097. URL: https://www.medrxiv.org/content/10.1101/2020.09.26.20200097v1.

[116] Pulkit Gupta et al. "Deregulation of translation due to post-transcriptional modification of rRNA explains why erm genes are inducible". In: *Nat Commun* 4.1 (June 10, 2013). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Antibacterial drug resistance;Post-translational modifications;Translation Subject_term_id: antibacterial-drug-resistance;post-translational-modifications;translation, p. 1984. ISSN: 2041-1723. DOI: 10.1038/ncomms2984. URL: http://www.nature.com/articles/ncomms2984.

[117] MacDougall Conan and Polk Ron E. "Antimicrobial Stewardship Programs in Health Care Systems". In: *Clinical Microbiology Reviews* 18.4 (Oct. 1, 2005). Publisher: American Society for Microbiology, pp. 638–656. DOI: 10.1128/CMR.18.4.638-656.2005. URL: http://journals.asm.org/doi/full/10.1128/CMR.18.4.638-656.2005.

[118] N. C. Gordon et al. "Whole-Genome Sequencing Reveals the Contribution of Long-Term Carriers in Staphylococcus aureus Outbreak Investigation". In: *J Clin Microbiol* 55.7 (July 2017), pp. 2188–2197. ISSN: 0095-1137. DOI: 10.1128/JCM.00363-17. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5483921/.

[119] Fabienne Antunes Ferreira et al. “Rapid nanopore-based DNA sequencing protocol of antibiotic-resistant bacteria for use in surveillance and outbreak investigation”. In: *Microbial Genomics* 7.4 (Apr. 22, 2021). ISSN: 2057-5858. DOI: 10.1099/mgen.0.000557. URL: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000557.

[120] W. Florian Fricke and David A. Rasko. “Bacterial genome sequencing in the clinic: bioinformatic challenges and solutions”. In: *Nat Rev Genet* 15.1 (Jan. 2014). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 1 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Bacterial genomics;Next-generation sequencing Subject_term_id: bacterial-genomics;next-generation-sequencing, pp. 49–55. ISSN: 1471-0064. DOI: 10.1038/nrg3624. URL: http://www.nature.com/articles/nrg3624.

[121] Jennifer L. Gardy and Nicholas J. Loman. “Towards a genomics-informed, real-time, global pathogen surveillance system”. In: *Nat Rev Genet* 19.1 (Jan. 2018), pp. 9–20. ISSN: 1471-0056, 1471-0064. DOI: 10.1038/nrg.2017.88. URL: http://www.nature.com/articles/nrg.2017.88.

[122] Joshua Quick et al. “Real-time, portable genome sequencing for Ebola surveillance”. In: *Nature* 530.7589 (Feb. 2016), pp. 228–232. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature16996. URL: http://www.nature.com/articles/nature16996.

[123] Nuno Rodrigues Faria et al. “Mobile real-time surveillance of Zika virus in Brazil”. In: *Genome Med* 8 (Sept. 29, 2016). ISSN: 1756-994X. DOI: 10.1186/s13073-016-0356-2. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5041528/.

[124] Narender Kumar et al. “Evaluation of a fully automated bioinformatics tool to predict antibiotic resistance from MRSA genomes”. In: *Journal of Antimicrobial Chemotherapy* 75.5 (May 1, 2020), pp. 1117–1122. ISSN: 0305-7453. DOI: 10.1093/jac/dkz570. URL: https://doi.org/10.1093/jac/dkz570.

[125] Louise J. Pankhurst et al. "Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study". In: *Lancet Respir Med* 4.1 (Jan. 2016), pp. 49–58. ISSN: 2213-2619. DOI: 10.1016/S2213-2600(15)00466-X.

[126] The CRyPTIC Consortium and the 100,000 Genomes Project. "Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing". In: *N Engl J Med* 379.15 (Oct. 11, 2018), pp. 1403–1415. ISSN: 0028-4793, 1533-4406. DOI: 10.1056/NEJMoa1800474. URL: http://www.nejm.org/doi/10.1056/NEJMoa1800474.

[127] Ahmed Babiker et al. "Use of online tools for antimicrobial resistance prediction by whole-genome sequencing in methicillin-resistant Staphylococcus aureus (MRSA) and vancomycin-resistant enterococci (VRE)". In: *J Glob Antimicrob Resist* 19 (Dec. 2019), pp. 136–143. ISSN: 2213-7173. DOI: 10.1016/j.jgar.2019.04.006.

[128] N. C. Gordon et al. "Prediction of Staphylococcus aureus Antimicrobial Resistance by Whole-Genome Sequencing". In: *Journal of Clinical Microbiology* 52.4 (Apr. 1, 2014), pp. 1182–1191. ISSN: 0095-1137. DOI: 10.1128/JCM.03117-13. URL: https://jcm.asm.org/content/52/4/1182.

[129] Maisem Laabei et al. "Predicting the virulence of MRSA from its genome sequence". In: *Genome Res.* 24.5 (May 1, 2014). Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 839–849. ISSN: 1088-9051, 1549-5469. DOI: 10.1101/gr.165415.113. URL: https://genome.cshlp.org/content/24/5/839.