2016

Assessment of Staphylococcus aureus Genetics: Clinical versus Community Epidemiology

Matthew Lawrance
University of Central Florida

ASSESSMENT OF *STAPHYLOCOCCUS AUREUS* GENETICS: CLINICAL VERSUS COMMUNITY EPIDEMIOLOGY

By

MATTHEW F. LAWRANCE
B.S. University of Central Florida, 2013

A thesis submitted in partial fulfillment of the requirements
for the degree in Master of Science
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2016

Major Professor: Christopher L. Parkinson

ABSTRACT

Staphylococcus aureus has an historical relationship with anthropogenic environments, particularly hospitals, where infection characteristics differ from community-acquired disease. This has promoted a designation of strains as healthcare or community associated. Despite this affiliation, genetic approaches have failed to support these groupings. In order to establish the genetic relationship between *S. aureus* from differing anthropogenic environments, I have analyzed the relatedness between three cohorts of *S. aureus*: nasal carriage isolates from community participants, infectious isolates from hospitals, and a cohort from an uninvestigated environment, an ambulatory clinic. Multilocus Sequence Typing (MLST) and *Staphylococcus aureus* protein a (*spa*) repeat regions were analyzed and the genetic relationships between cohorts at these sites were determined. I found high similarity in recovered sequences within and between all cohorts, with cohorts sharing 100% sequence identity across some samples. Phylogenetic reconstruction of the combined datasets indicate panmixia, with samples of all origins belonging to shared genetic lineages. Additional clustering algorithms supported this pattern. The findings of this study indicate that there is strong genetic similarity between both infectious strains and nasal carriage strains and between isolates from all cohorts. This research has implications for healthcare, as it demonstrates that *S. aureus* from differing environments are genetically similar (often identical), cautioning against delineating strains into nasal carriage or infectious based on origin. This research also informs the study of *S. aureus* evolution – strengthening the conclusion that differentiation at stably selected markers in lineages within

differing ‘healthcare habitats’ is insufficient to explain observed phenotypic differences, and alternative explanations must be explored.

ACKNOWLEDGMENTS

I would first like to extend my gratitude towards my committee members, Doctors Alexander Cole, and Anna Savage for their enormous patience and willingness to inform me during the formation of my thesis. Without their varied expertise in the areas of pathogen epidemiology and evolutionary analysis as a whole, the completion of this work would have been impossible. The comments and advising that I received both on this document and the manuscripts authored as a result of work relating to this thesis were beyond helpful and contributed overwhelmingly to the successful completion of this work.

I would like to sincerely thank Doctor Christopher L. Parkinson, my major advisor, whose willingness to adopt me into his lab was a major indicator of his faith for which I will always be humbled. Without his tutelage and encouragement, I firmly believe that I would have not only failed to advance my thesis work to completion, but I would have also failed to advance myself as an academic, and failed to advance my broader conceptualization of evolution and biology as a whole.

I would also like to acknowledge the contribution of my fellow graduate, undergraduate and faculty members of the PHASED lab group. All individuals in this group served as a font of constructive criticism and provided varied perspectives in interpreting the content of this work. Particularly, I acknowledge Doctors Eric Hoffman and Tiffany Doan for their astute and much-needed input, both in the interpretation and presentation of my thesis work. Without their critical and accurate assessments, and the input of the PHASED group overall, the quality of the work herein would not have been achieved.

The contribution of many individuals were required in order to coordinate sample collection. In particular, I acknowledge Doctors Michael Deichen and James Schaus for allowing access to the University of Central Florida Health Services to facilitate sampling. I also acknowledge the interest of both Doctors in being involved in the investigative process that led to the contents of this thesis. I would also like to acknowledge John Deichen, who provided invaluable assistance in acquisition of samples.

This project would not have been possible without the work of the Cole Lab. I thank Doctor Ryan Lamers for providing the context of a UCF community sampling scheme as well as the samples themselves for comparison. I thank Doctor Gowrishankar Muthrukrishnan for his direct assistance in handling both the samples utilized in this thesis as well as the resulting data.

I recognize the Parkinson lab, particularly Jason Strickland, Jason Hickson, Andrew Mason, Rhett Rautsaw and Katie Mercier, for providing repeated and positive suggestions for the direction and presentation of this research, as well as their support and camaraderie.

Finally, I thank my friends and family for their support and patience. I imagine the latter miss seeing me something fierce.

TABLE OF CONTENTS

LIST OF FIGURES vii
LIST OF TABLES viii
LIST OF NOMENCLATURE ix
LIST OF ACRONYMS x
BACKGROUND 1
INTRODUCTION 8
MATERIALS AND METHODS 12
Ethics Statement 12
Bacterial Isolates 12
DNA Extraction 13
DNA Sequence Analysis 13
Phylogenetic Reconstruction 14
Population Structure: MLST Data 16
RESULTS 17
Multilocus sequence typing demonstrates similarity between isolates of varying origins 17
Phylogenetic reconstruction of MLST data shows lack of differentiation between taxa 18
eBURST produces clusters containing *S. aureus* from all sample categories 19
DISCUSSION 21
CONCLUSION 27
APPENDIX A: TABLES 29
APPENDIX B: FIGURES 34
REFERENCES 40

LIST OF FIGURES

Figure 1: Schematic representation of all competing hypotheses 35
Figure 2: Logistic Regression of donor characteristics as predictors for carriage. 36
Figure 3: Most likely phylogenetic tree. 37
Figure 4: MLST EBURST of samples. 38
Figure 5: Minimal Spanning network of MLST data. 39

LIST OF TABLES

Table 1: Demographic information of participants (a) and positive donors (b) by race 30
Table 2: Sequence of Primers utilized in PCR 31
Table 3: Partition of gene fragments and codon positions..... 32
Table 4: Bayesian Factor Hypotheses testing of various competing hypotheses..... 33

LIST OF NOMENCLATURE

AMBULATORY CARE CLINIC	A healthcare center that provides professional medical care on a strictly outpatient basis
CARRIAGE	Host status in which a potentially infectious agent is present, but incurring no symptoms or damage to host processes, *see COLONIZATION*
COHORT	A group of samples collected due to a shared characteristic (i.e., sample location, time of sampling, ect.)
COLONIZATION	Infectious agent status in which agent resides on or within a host without incurring consequences to the host, *see CARRIAGE*
DISEASE	An abnormal condition of an organism limiting normal function; the result of the interaction of infection and host response, or abnormal host response in the absence of infection
INFECTION	Invasion and multiplication of an agent within a host such that the host develops clinical symptoms due to cellular or systemic damage resulting from the actions of that agent
NOSOCOMIAL	Pertaining or originating in a hospital

LIST OF ACRONYMS

ACC	Ambulatory Care Centers
BF	Bayes Factor
BI	Bayesian Inference
CA	Community-associated
HA	Healthcare-associated
MLST	Multi-Locus Sequence Typing
PVL	Panton Valentine Leucocidin
SCCMec	Staphylococcus Cassette Chromosome mec
Spa	Staphylococcal Protein A
TLR2	Toll-like Receptor 2

BACKGROUND

Staphylococcus aureus was first described by Sir Alexander Ogston in 1880, when it was discovered in the pus of a surgical incision (Ogston 1882). By 1884, it was defined as the type species for the genus *Staphylococcus*. A gram positive, coagulase positive coccus, *S. aureus* is among the most prevalent species in its genus (alongside *Staphylococcus epidemidis*), and one of the most common constituents of human microflora, as well as being a prominent colonist of livestock and other species (Tancrede 1992). As a commensal organism, *S. aureus* primarily colonizes mucosal membranes, including the throat, the epithelium of the anterior nares, within the vaginal mucosa, and on the exterior skin, particularly in the axillary and groin (Payne 1966). Colonization of the human host is facilitated by a number of adhesion-receptor interactions, where bacterial tethering proteins bind to carbohydrate moiety in mucin (Shuter 1996). Asymptomatic colonization of *S. aureus* is of two types: persistent carriers acquire a single strain of *S. aureus* which they carry for extended periods of time. Longitudinal sampling of these individuals demonstrate the persistence of these strains (Wertheim 2005). Intermittent carriers alternate between colonization, clearance, and recolonization (Williams 1963). Variance in host factors, particularly toll-like receptor 2 (TLR2), are strongly implicated in colonization success (Lorenz 2000). Asymptomatically carrying *S. aureus* longitudinally – especially persistently – is a strong indicator of eventual pathogenesis (Kluytmans 1997). *Staphylococcus aureus* physiology strongly promotes opportunistic invasion of other tissues, where infection and disease progression occurs. Collagen binding proteins facilitate invasion of exposed skeletal tissue, particularly joints, while clumping factors and fibrinogen binding

factors allow for *S. aureus* to colonize incisions on the epithelial tissue (Rohde 2007, Eidhin 1998). When the circulatory system is accessible, *S. aureus* tethering proteins are capable of inducing endothelium tissue endocytosis, allowing escape from host immune defenses and infection of the blood (Clark 2009).

Widespread use of antibiotics has led to a *S. aureus* developing numerous resistances, altering this species' epidemiology. With the rare exception, such as vancomycin, *S. aureus* evolves antibacterial resistance rapidly, typically demonstrating tolerance of new antibiotics within the span of a few years (Klevens 2007, CDC 2002, Tsiodras 2001). While drugs such as vancomycin maintained efficacy against *S. aureus* for several decades, antibiotic resistance to even these treatments have been shown (Tabaqchali 1997). In the case of penicillin, *S. aureus* demonstrated resistance prior to the widespread initial use of the drug (Rammelkamp 1942). *Staphylococcus aureus'* rapid acquisition of resistance is owed in part to Chromosome Recombinase, an enzyme capable of excising regions of *S. aureus'* genome (termed chromosome cassettes) and exporting them in horizontal transfer events, allowing unrelated lineages to rapidly adopt new resistances. *Staphylococcus aureus'* ability to horizontally transfer these factors is additionally accelerated by the presence of biofilms – a deposited extracellular matrix this species is known for (Rohde 2007). Of particular concern for healthcare, methicillin resistant *S. aureus* (MRSA) acquires resistance via the exchange of *Staphylococcal* Cassette Chromosome mec (SCCmec, Ito 2004, Katayama 2000). MRSA strains are more aggressively infectious, harder to manage, and more costly to treat than their methicillin susceptible (MSSA)

counterparts (Capitano 2003, Chambers 2001). The rapid acquisition of resistance to broad spectrum antibiotics prompted the Centers for Disease Control (CDC) to label *S. aureus*, particularly MRSA strains, as high concern for human health. Over 500,000 hospitalizations occur annually in the United States, incurring a mean economic cost of 1,332 USD (MSSA) to 2,607 USD (MRSA) (Capitano 2003).

Given the threat *S. aureus* poses to human health, extensive efforts have been made to clarify its diversity, resulting in a suite of molecular techniques aimed at characterizing strains. Pulse Field Gel Electrophoresis (PFGE) is the process of running gel chromatography on whole-protein samples with alternating current angle and direction to achieve maximum separation of protein bands by weight (Kaufmann 1998). PFGE remains the method that results in the highest resolution between strains, allowing for the discrimination between samples that cannot be differentiated through other methods. However, this method is sensitive to variation in technique and conditions between labs, meaning comparing results across studies is not possible (Murchan 2003). Multi Locus Sequence Typing (MLST) was designed to address this issue. MLST is applicable to a broad range of microbial species, and requires the sequencing of several housekeeping loci. In the case of *S. aureus*, this corresponds to seven housekeeping loci (*arcC, aroE, glpF, gmk, pta, tpi, yqiL*). Newly sequenced samples are compared to a curated online database composed of all previously known alleles, and assigned known numerical labels in the case of a match. Novel alleles are given new labels and curated in the database. Unlike PFGE, MLST is broadly applicable across studies, at the cost of marginal reduction in

discriminatory power between strains. However, as MLST requires the use of housekeeping genes, which by nature slowly accumulate variation due to negative selection, the ability to discern closely related samples at local scales suffers (Enright, 2000). New markers have been developed in response, including clumping factors a/b, fibrinogen binding factors, and *Staphylococcal* protein a (*spa*). Virulence factors such as these are far more variable than housekeeping genes, and enhance discriminatory power of investigations while simultaneously allowing results to be shared across labs.

The *spa* locus, specifically, codes for a membrane-tethered virulence factor involved in host-immune evasion. The portion of the protein extending extracellularly interacts with the Fragment Crystallizable (Fc) region of human Immunoglobulin G (IgG), preventing native IgG activation of effector immune cells. The membrane spanning region of this protein is of interest for evolutionary analysis, as it is formed from variable repeat regions in the corresponding portion of the *spa* locus. The presence and absence of certain nucleotide repeat regions, as well as the order in which those repeats occur, allow homology statements to be drawn between compared isolates. Typing of strains based on sequence identity match at these repeat regions results in discriminatory power of 98% when compared to PFGE (Harmsen 2003).

Investigations utilizing these techniques have revealed broad themes in *S. aureus* genetic structure. Notably, *S. aureus* is a remarkably clonal organism. Recombination does occur, but at very reduced rate when compared to single-nucleotide point mutations, which occur 15 times more frequently (Feil 2003). As a result, *S. aureus* isolates tend to cluster into clonal complexes

(CCs), where a single strain founds an initial population, and subsequent strains originate via point mutations from the founder. Resultantly, large demographics of *S. aureus* are composed of only a few clonal complexes. For example, all MRSA isolates in the Meuse-Rhine Region are of four European clonal complexes (Deurenberg 2005).

Given the larger clonal genetic structure of *S. aureus*, it is expected that a lack of inherited variation across lineages would result. However, *S. aureus* does show remarkable phenotypic differentiation across environments. In particular, *S. aureus* varies across healthcare environments, both in specific proteins and in infection profile. This scenario has led to the adoption of a community- or hospital-associated designation of strains (Lodise 2003, Steinberg 1996, Vandenesch 2003). In hospitals, *S. aureus* antibacterial resistance occurs at significantly higher rates than within the community at large (Panlilio 1992, David 2010). Additionally, hospital-associated strains are responsible for scalded-skin syndrome in infants in neonatal wards, and terminal bacteremia infections (Ladhani 1998, Cosgrove 2013). Within hospitals, it is hypothesized that transmission of *S. aureus* is facilitated by direct attendant transmission, accelerating infection rates and therefore strain virulence (Ewald 1993). In contrast, community-associated strains exhibit lower prevalence of antibacterial resistance, and terminal illness generally results from necrotizing pneumonia – an aggressive condition that usually killed within 72 hours of onset (Gillet 2007). These phenotypic differences are the result of underlying proteome variation. Hospital-associated *S. aureus* are more likely to acquire methicillin resistance via large SCCmec types I, II, or III. Additionally, they are more likely to

possess exfoliatin, an exotoxin responsible for scalded skin syndrome (Zhang 2005 , Melish 1972). Methicillin resistant Community-associated *S. aureus* possess smaller SCC mec type IV and V, in addition to Panton-Valentine Leucocidin (PVL). PVL is responsible for invasion and destruction of lung tissue by those strains (Xu Ma 2002, Vandenesch 2003).

Both SCC mec and PVL are elements that are preferentially found in either community or hospital environments as significantly higher rates than the opposing habitat. However, *S. aureus'* genome can be broadly divided into three categories: core genes, core variables, and mobile genetic elements (MGEs). MGE regions include portions of the genome encoded by prophages and *Staphylococcal* cassettes. Virulence and antibiotic resistance loci, including those that are associated with healthcare environments, are disproportionately represented within MGE regions (Lindsay 2010). In addition to virulence and antibiotic resistance proteins, many cassettes also encoded for the enzymes needed for their horizontal transmission (recombinases), accelerating their propagation independent of *S. aureus* reproduction. As the phylogenetic signal originating from MGE regions may not reflect the evolutionary history of *S. aureus* overall, investigations have been standardized on core variable loci, which include previously mentioned MLST and *spa* loci. Attempts to recover any signal of variation across healthcare environments utilizing core variables have been unable to resolve hospital- or community-associated strains into monophyletic clades (Lamers 2011). However, portions of the core variable regions, including *spa* have not been utilized in phylogenetic reconstruction. Given the lack of clarity surrounding the relationships of phenotypic variation and genetic

differentiation within this species, it is important to increase taxonomic sampling, particularly by investigating new loci and sampling locations.

INTRODUCTION

Staphylococcus aureus is a widespread human pathogen of high concern to global human health (Noskin 2007). This pathogen is most often associated with healthcare settings, where it is the leading cause of nosocomial infections (Grundmann 2006). In addition, *S. aureus* is commonly found affecting community demographics not associated with hospital exposure, where asymptomatic carriage is implicated in eventual infection (van Belkum 2009). An estimated 20-30% of the global human population is colonized by *S. aureus* persistently, with intermittent carriers ranging from 60%-100% of sampled individuals in given populations (Heiman 2005, Williams 1963, Dancer 1991, Nouwen 2004, Van Belkum 2009). Despite being found commonly in both clinical and nonclinical environments, the narrative in *S. aureus* literature delineates strains into one of two categories: hospital-acquired (HA) strains, and strains that are community acquired (CA) (Karauzum 2008, DeLeo 2010, Cheung 2011). Phenotypic variation in pathogenesis factors of strains and the symptoms of resulting infections are suggestive of such categories (Rudd 2008). Despite a strong historic focus on hospital samples, there is growing interest into the epidemiological consequences of community carriage, reemphasizing the HA/CA divide (Muthukrishnan 2013).

Clarifying relationships in species of clinical concern is important in explaining the epidemiology of that species (Monis 1998, Oliver 1996). Recent research indicates that genetic structure of *Staphylococcus* is poorly understood and in need of revision (Lamers 2012). Systematic investigations into the underlying genetics of HA and CA isolates of *S. aureus* have failed to recover evidence to differentiate these categories (Lamers 2011). This conclusion is surprising,

given known protein variation between HA and CA strains of *S. aureus* (Montgomery 2010, Deurenberg 2008). In order to address this disconnect, further genetic investigations of *S. aureus* are necessary, particularly those sampling more variable markers (>1% mean pairwise distance, Cooper 2006).

Classically, samples of *S. aureus* representing the HA designation have been drawn from hospital settings (Holden 2004). However, recent decades have seen the emergence of new healthcare facilities distinct from both hospital and community environments. Medical clinics, also known as 'walk-in clinics' or ambulatory care centers, share similarities between both traditional hospital settings as well as nonclinical, community environments (Starfield 1991). These similarities, including attendant mediated care and minor surgical procedures, are known to have consequences for pathogenesis and disease (Ewald 1993, Kluytmans 1997). Given their distinct nature, usage and demographic data pertaining to ambulatory care centers are generally collected independently of hospitals (Schappert 1998). Despite this acknowledgment of ambulatory care centers as unique environments, little research has been conducted into the pathogens circulating within them. Though investigations into the epidemiology of *S. aureus* in ambulatory care centers have been performed, no information regarding isolate genetic relationships among *S. aureus* isolates within these centers is available (Szumowski 2007). Ambulatory care clinics may influence the relationships within *S. aureus,* given that individuals – and presumably the pathogens they carry – permeate through these environments more rapidly than hospitals, while frequent direct contact by primary care providers within these clinics more closely resembles hospital procedures. Population genetics investigations including

isolates from ambulatory care centers are necessary for shedding light into the epidemiology of *S. aureus* in this previously unexplored environment.

S. aureus population structure and phylogenetic relationships have been extensively characterized. Pulse-field gel electrophoresis (PFGE) and multi-locus sequence typing (MLST) are methods most often utilized in these investigations (McDougal 2003, Enright 2000). While PFGE results in the highest levels of resolution and differentiation between strains, the method is difficult to reproduce precisely, undermining results shared between laboratories and investigations (Chung 2000). Comparatively, MLST marginally sacrifices resolution between samples to achieve standardized results that are easily compared between studies (Maiden 1998, Enright 2002). However, the loci utilized in MLST are housekeeping genes, the functions of which are vital in preserving *S. aureus'* viability. Thus, these loci are, by nature, slowly evolving, and the ability to distinguish between closely related samples (e.g., those collected from closely-associated localities) suffers (Enright 1999). In order to address this, standardized methods aimed at factors exhibiting high variability even at local scales (e.g., virulence factors) have been developed. Staphylococcal protein A (*spa*) typing has emerged as a compromise between the resolution of PFGE and the applicability of MLST (Harmsen 2003). While *spa* typing and MLST have been used in the thorough investigation of *S. aureus* population structure within the community and hospitals, no investigation has yet sought to leverage these tools in exploring novel environments such as ambulatory care centers.

Here, I have performed an evolutionary analysis of seven MLST gene fragments and *spa,* incorporating samples taken from a representative example of previously uninvestigated ambulatory care center, with the aim of better understanding how population structure of *S. aureus* varies across heterogeneous environments. Isolates were taken from employees of the University of Central Florida's Health Center, a representative ambulatory care center. This sample site was additionally attractive as it was adjacent to a previous sampled community, allowing us to eliminate the effects of geographic distance which can influence the resulting phylogeny (Lamers 2011, Muthrukrishnan 2012, Slatkin 1990). Given the systematic uncertainty within *S. aureus* and the specific inability to validate differences between hospital and community isolates, I hypothesized that isolates from the ambulatory care center will not be genetically differentiated based on their origin. Therefore, I predicted similar sequence identify, shared haplotypes, and no signal of differentiation within the phylogenetic reconstruction among and between all cohorts.

MATERIALS AND METHODS

Ethics Statement

Samples were collected from willing participants, under the guidance of procedures approved by University of Central Florida's Institutional Review Board (IRB) approved procedures (Muthrukishnan, 2012). Informed consent was acquired from all participants prior to sampling. All investigators involved in sample collection were properly instructed and granted Collaborative Institutional Training Initiative (CITI) certification.

Bacterial Isolates

141 healthy employees of the University of Central Florida's Health Center underwent pre-screening for bacterial isolates. Participants volunteered demographic information in order to inform the population structure of recovered bacterial isolates. Of the screened participants, 29 (20.5%) resulted in positive identification of *S. aureus*. Isolates were collected via participant insertion of a cotton swab into both nostrils and circulation for approximately 5-10 seconds. Swabs were immersed in glycerol-Trypticase™ Soy Broth (TSB) solution during transport, followed by plating on Trypticase™ Soy Agar (TSA) impregnated by 5% sheep's blood. Isolates were incubated at 37^{o}C for 16 hours. Resultant colonies were tested with Staphyloslide™ Latex Test reagent to positively identify cultures as *S. aureus*. Verified *S. aureus* colonies were isolated and inoculated in TSB for an additional 16 hours at 37^{o}C at 250rpm in preparation for

DNA extraction. Chi-squared tests of donor carriage were performed to determine the uniformity of carriage within the ACC cohort.

DNA Extraction

1.5mL of each bacterial inoculate was centrifuged at 16,000*g* for two minutes. Supernatant was removed, and the remaining pellet was utilized in the extraction protocol. DNA was extracted utilizing GenElute Bacterial Genomic DNA kits (Sigma-Aldrich Co., St. Louis, Missouri), in accordance with manufacturer's instructions. Fragments of seven MLST loci (*arc, aroE, glpF, gmk, pta, tpi, yqiL)* ranging from 402-512 base pairs were amplified (Enright 2000). Additionally, approximately 500 base pair fragments of *spa* were amplified. Approximately 30 ng of genomic DNA was added to 30 uL reactions mixtures containing: .02U/uL Platinum™Taq DNA Polymerase high Fidelity, 1X PCR buffer, 2mM $MgSO_4$, 0.3 mM dNTPs, 0.3 uM of external forward and reverse primers (Table 1), and 2% dimethyl sulfoxide (DMSO). PCR amplification was achieved using Peltier Thermal Cyclers (PTC) 200 (MJ Research) for 30 cycles under the following cycle parameters: initial denaturation at 95^{o}C for five minutes, annealing at 55^{o}C for one minute, extension at 72^{o}C for one minute, followed by denaturation at 95^{o}C for one minute. A final extension step at 72^{o}C for five minutes followed.

DNA Sequence Analysis

Following amplification, PCR products were purified utilizing QiaQuick PCR Purification Kit (QiaGen, Redwood City, Ca). All Sanger Sequencing was performed at University of Arizona's Genetics Core. Forward and reverse reads were visualized with Sequencher 5.1 (Gene Codes

Co., Ann Harbor, Michigan). Sequences were organized in MEGA 5.2 and aligned with ClustalW. Sequence Types (STs) were determined for each sample based on alleles identified for each of the seven MLST loci. Alleles were cross-referenced against the *S. aureus* database curated at MLST.net. Novel alleles – or novel combinations of known alleles – were submitted to the MLST database for curation, whereupon new allele designation and STs were obtained.

Phylogenetic Reconstruction

In order to infer the relationships between samples of hospital, community, and ambulatory care center origins, phylogenetic analysis was performed on a concatenated dataset of all eight sequenced loci for all samples. Datasets representing community samples were constructed utilizing 141 samples from a previous study of the University of Central Florida's student, staff, and faculty body. Examination of this data revealed no overlap from samples from the community and clinical cohorts. Gene sequences from 15 strains associated with hospital infection were acquired from NCBI GenBank and included in this analysis. The previously sequenced hospital strains were N315, Mu50, COL, MRSA252, MSSA476, MW2, USA300_FPR3757, NCTC8325, JH1, JH9, Newman, Mu3, USA300_TCH1516, 04-02981, and TW20. These strains were selected as they were both utilized as references in the previous UCF community sampling (Lamers, 2011) as well as being representative of common clonal types found within hospitals across the United States. The concatenated dataset of all samples was partitioned by gene fragment and codon position, with models of evolution being assigned by Akaike Information Criterion (AIC) within PartitionFinderv1.1.1 (Lanfear 2012) (Table 3). Once partitioned, phylogenetic reconstruction of the data was performed using Metropolis-Hastings

Coupled Markovnikov Chain Monte Carlo in BEAST v1.8 (Drummond 2012). Two independent Bayesian Inference (BI) runs were carried out using random starting trees. After five million iterations, runs were terminated and visualized utilizing Tracer v1.5 (Rambaut 2014). The most-likely tree was summed utilizing TreeAnnotatorv1.8.2 and visualized with FigTreev1.4.0 (http://tree.bio.ed.ac.uk/software/figtree).

In order to test the likelihood of competing hypotheses against the unconstrained tree, Bayes Factor Topography Testing was performed (Figure 1). Additional BI runs were performed, utilizing identical operational parameters (run length, number of chains, ect). BEAST v1.8 allows for the forced monophyly of designated taxa. Summarized phylogenies were reconstructed for my dataset when all taxa were constrained to be monophyletic given their origins (i.e., HA, AAC, or CA). The log marginal likelihood of the unconstrained, maximum-likelihood tree (B_0) and the constrained hypotheses trees (B_A) were calculated using two methodologies: Bayesian Stepping Stone and Path Sampling. Bayes Factors for alternative hypotheses were calculated as $BF = \frac{P(D|M_0)}{P(D|M_A)}$ where $P(D|M_x)$ is the marginal likelihood of model x – essentially a ratio between marginal likelihoods of competing models. Calculations for Bayes Factors were adopted from Rafferty, 1996. Interpretation of Bayes Factors were informed by Posada and Buckley 2004, stating that Bayes Factors exceeding 150 are considered very strong indication of a less likely tree, with those falling between 12 and 150 considered good indication of a less likely tree, and those between 3 and 12 weakly indicative of a less likely tree. Bayes Factors below 3 are considered insignificant.

Population Structure: MLST Data

Sequence Types were grouped based on clusters sharing six of seven MLST alleles in common with at least one other member of the cluster. Clustering was performed in eBURSTv3 (Feil 2004), which provides for a query set of clinical samples to be tested against the reference set of community and hospital samples. Clustering of samples of six of seven shared alleles is reflective of eBurst procedures designed to account for the potential for high variation in single alleles to result from recombination, rather than differentiation within the lineage. Bootstrap support for clusters were inferred based on 1,000 resampled iterations.

In order to visualize population structure at MLST loci without masking recombination at a single locus, PopART (http://popart.otago.ac.nz) was utilized to produce a minimum spanning network of all mutational steps.

RESULTS

Multilocus sequence typing demonstrates similarity between isolates of varying origins

MLST analysis of the 29 samples isolated from my sampling scheme within UCF Health Services identified 10 different sequence types (STs), one of which was new. Of these 10 STs, five were previously recovered in a sampling of the adjacent community. Of the remaining five samples, four were previously identified as HA strains (STs 105,109, 20, and 1) and one was novel. Overall, carriage rate within the ACC cohort was 20.6%. χ^2 tests of demographic information indicated that carriage was not evenly distributed across groups. Notably, carriage was 10% for non-Hispanic black donors. This difference was predominately driven by a low carriage rate of 4.5% by non-Hispanic black females. Non-Hispanic black males had a carriage rate of 25%. Asian/Pacific Islanders had a significantly higher than average carriage rate at 36%. This result was driven by females within this group (50%). No tested male Asian/Pacific Islander was positive for *S. aureus* carriage. Neither age of donor nor length of employment were significant predictors of carriage, though length of employment did approach significance (p=.07, figure 2). It appears that there is an inverse correlation between extensive employment within the health center and likelihood of carriage, possibly indicating an increase in hygienic practices and caution in inveterate primary care physicians.

Of the 10 recovered sequence types, the most prevalent was ST30, accounting for 9 of 29 isolates (31%). Sequence types 5 and 45 were the second and third most commonly identified sequence types following ST30, with 6 (21%) and 4 (14%) instances respectively. These results

are roughly comparable with the prevalence of prominent sequence types in the prior UCF sampling (Lamers, 2011).

Phylogenetic reconstruction of MLST data shows lack of differentiation between taxa

Genetic similarities of ACC isolated were apparent from the most likely phylogenetic tree (Figure 2). In the unconstrained tree, no monophyly between samples of clinic, hospital, or community origin were recovered; clades containing combinations of all three origins were recovered with strong nodal support. Investigation into the demographic information of the donors of these samples revealed that they were in no shared demographic category (dissimilar ages, sex, race, position, ect). Many samples in the phylogeny were grouped into large, unresolved polytomies containing HA, CA and ACC isolates, indicating low differentiation amongst those samples consistent with recent common ancestors and ongoing geneflow between origins.

In order to determine the strength of the most likely tree compared to those that represent competing hypotheses, additional trees were constructed, demarcated by factors that could potentially influence evolutionary trajectories (locality of sampling, ethnicity of participant, profession of participant, ect). All trees representing competing hypotheses were significantly worse at representing the data than the most likely unconstrained tree, based on BF analysis (Table 3).

I was also interested in determining the relative contribution of genetic markers of varied origin (ie, housekeeping MLST vs variable repeat regions in *spa*) on my reconstruction. Additional phylogenies were constructed, weighting the conserved genes more heavily than the hypervariable *spa* locus. Weights were assigned based on mean pairwise genetic distance between samples. Average genetic distance between all seven MLST loci was .6%; between *spa* alleles, comparatively, there was a mean genetic distance of 4.7%. The marginal likelihood of the resulting weighted phylogeny was compared against the unconstrained tree (Table 4). The weighted tree vastly outperformed the unweighted tree in my Bayes Factors hypothesis-testing framework (BF>150).

eBURST produces clusters containing *S. aureus* from all sample categories

MLST eBURST clustered samples from all origins (ACC, HA, CA) into 10 distinct groups, with 13 remaining singletons that were not assigned to a cluster (Figure 3). These singleton STs were most often represented by a single isolate, though ST 20 was not clustered, and contained two individual isolates from my clinical subset. Of my clusters, three (3, 7, 9) contained only community samples. All other clusters contained an admixture of all categories of samples. Groups 2 and 8 were founded by CA STs, with the remaining clusters founded by shared sequence types.

In order to demonstrate inferred haplotypes and interconnect clusters without the influence of hypothetical recombination, PopArt was utilized on MLST data in order to build a minimum spanning network (Figure 4).The results acquired from the minimum spanning network were in

agreement with the clusters resolved by eBURST. Highly successful haplotypes (e.g., ST 30, ST 5) were shared between both ACC and CA isolates, with no monophyly of samples for any sample origin.

DISCUSSION

Well characterized relationships with few uncertainties is important when examining the epidemiology of pathogenic species. Due to the nomenclature surrounding *S. aureus* relationships across healthcare environments, it was the goal of this study to examine the evolutionary trajectory, relatedness and genetic structure of *S. aureus* in a previously unstudied medical environment in comparison to previously assessed cohorts. As STs originating in my ACC cohort were similar, and often identical, to CA and HA samples, I conclude that ACC do not represent a distinct environment containing independently evolving *S. aureus*. The inclusion of a new molecular marker of high sequence variation within my phylogenetic reconstruction was expected to increase resolution when compared to previous reconstructions. I never-the-less found no major clade exclusive to a single sample origin, strengthening previously drawn conclusions from phylogenetic reconstructions of only HA and CA isolates with only MLST data (Lamers 2011). Ultimately, I conclude that no significant variation divides *S. aureus* samples in regards to the origin of isolation included in this study.

The potential divide between community and healthcare strains may seem to impact the effectiveness of sampling schemes such as the one employed in this study. On first examination of *S. aureus* literature, HA *S. aureus* appears adapted to causing disease, as representative examples of healthcare associated strains including those included in my phylogenetic reconstruction are almost universally isolated from infected, not asymptomatic, individuals (Tenover 2004, Tsiodras 2001, Rotun 1999, Cosgrove 2013). It could be argued that by sampling only healthy individuals within my cohort and the reference community cohort, my scheme

biased my search against *S. aureus* strains adapted for disease, and therefore healthcare associated strains. However, the relationship between HA *S. aureus* and disease progression is likely spurious and the reflective of the treatment demands of hospitals. This is not surprising, given that the primary interest of hospital case studies is the monitoring and description of disease; individuals not exhibiting disease, including those asymptomatically colonized, warrant no case study. In short, most HA samples have come from individuals who first became diseased, and then subsequently were sampled without having been screened for the same carriage asymptomatically prior (Tenover 2004, Tsiodras 2001, Rotun 1999, Cosgrove 2013). Studies that have sampled hospitalized individuals for asymptomatic nasal carriage have shown that the same HA strain of *S. aureus* that ultimately results in fatal disease was first carried asymptomatically for some time before progressing to disease (Young 2012). Additionally, monitoring of asymptomatic carriage of new patients entering hospitals demonstrated that only a minority of patients newly colonized by *S. aureus* while hospitalized – the literal definition of HA – progressed to a diseased state (Davis 2004). Additionally, there is a growing body of evidence that demonstrates that community-associated strains are fully capable of progressing to infection and disease (Davis 2007, Rieg 2005, Voyich 2006). Ultimately, it appears that host-factors, rather than pathogen-adaptation, are responsible for disease progression (Cole 1999, Quinn 2007). In short, HA and CA designations do not relate to infectious and noninfectious strains, respectively, strengthening my confidence in the validity of my sampling.

Phylogenetic analyses perform best when moderately variable sequences are utilized to build homology statements between pairs of taxa. In instances such as this study, where genetic differentiation is very low, posterior support values tend to suffer. This lack of variation in MLST loci is attributed to their role as housekeeping loci that encode for proteins necessary for *S. aureus* viability, most commonly metabolic enzymes. In my cohort, the average genetic distance between MLST loci was less than a single percentage point (.6%). This tendency has been noted in previous phylogenetic reconstructions incorporating these loci, including the study that generated my reference community cohorts (Muthrukrishnan 2012). The lack of variation in these loci and the resulting potential loss of confidence in analyses in studies at the local scale has been known since the inception of these sites as markers (Enright 2000). Efforts have been made to design markers that retain MLST's ability to be exported between labs without sacrificing variability, including *spa*, clumping factors a/b, and fibrinogen binding factor (Harmsen 2003, Lamers 2011). My inclusion of *spa* in this study is the first phylogenetic reconstruction to utilize this region, and doing so increased the variability across my cohort, as *spa* demonstrated a mean genetic distance of 4.7%. I none-the-less recovered low posterior values at the tips of my tree, resulting in broad polytomies. This variation in my markers also demonstrated itself in the Bayes Factors I recovered for my weighted tree. Based on mean genetic distances across my loci of interest, it appears that MLST loci are eight times less likely to demonstrate nucleotide diversity between any two individuals at any given position, likely due to the influence of negative selection maintaining relative homogeneity at these sites (Enright, 2000). By informing the reconstruction to weigh autapomorphies at these loci eight

times more heavily than those in spa, I recovered a more likely tree than the unweighted phylogeny alone. This tree mainly increased posterior probabilities at deep nodes, however, and the polytomies at the tips of the tree were unresolved. This underscores the highly conserved nature of MLST loci, and demonstrates the need for new variable marker discovery in order to strengthen investigations into *S. aureus.*

These conclusions were largely confirmed by eBURST analysis. The purpose of clustering algorithms such as eBURST is to delineate groupings of samples into familial groups (in this case, clonal clusters). The clustering of our samples were consistent with a population reproducing across environments. As gene flow is high, with identical haplotypes being seen in both ACC and CA samples, it is unlikely that any lineage specific differentiation or adaptation in the core genome of *S. aureus* is likely to arise due to any divide caused by anthropogenic healthcare environments. Additionally, as the sequence data I utilized in these analyses are reflective of different regions of the core variable portions of *S. aureus* genome, it can be inferred from our results that historical differentiation has not occurred within these regions. Differentiation may be better explained by the phylogenetic reconstruction of alternative regions of *S. aureus'* genome – such as mobile genetic elements – which are known to harbor loci coding for specific proteins responsible for phenotypic variation in this species.

Ultimately, phylogenetic analyses, eBURST clustering, and minimum spanning haplotype networking of my samples all support the conclusion that isolates taken from ambulatory care clinic employees are not significantly isolated from isolates extracted both from healthy

community carriers or hospitalized individuals. This additionally supports the previous findings which assert that nasal carriage community strains were not themselves differentiated from pathogenic hospital isolates (Feil 2003, Wertheim 2005, Feil 2004) and underscores a lack of differentiation between groupings of *S. aureus* based on locality of sampling.

My research reinforces the growing confidence that *S. aureus* lineages are homogenous at core variable markers irrespective of sample origins, even when utilizing highly variable virulence loci such as *spa*. Additionally, as even geographically correlated samples (UCF ACC and UCF CA) were rendered paraphyletic and in polytomies with diffusely distributed HA samples, this genetic homogeneity in core variables is not seemingly influenced by the effects of distance. Based on this conclusion, future investigations into *S. aureus* differentiation would best be served by looking into variation into markers selected for variability – such as those coding for highly selected proteins that regulate the expression of phenotypes that vary between hospitals and the community. Additionally, the findings of my study clear the way for investigations into alternate explanations for apparent strain differentiation, such as phenotypic plasticity, or modulation of *S. aureus* biology based on host-response rather than pathogen-genetics. From a healthcare perspective, the conclusions laid forth by my study have implications for management of *S. aureus*. The failure to delineate strains based on origin of isolation would suggest the potential epidemiology of this species to be unified across environments, as there seems to be no genetic bases for differentiation. More specifically, these results caution against disregarding asymptomatically carried samples within ambulatory care clinics as potentially disease causing agents within those environments, as they are genetically indistinguishable

from strains proven capable of causing disease. Ultimately of concern from a healthcare perspective, the general homogeneity of *S. aureus* underlying genetics encourages the practice of assuming all strains have the potential to be infectious given the correct circumstances, regardless of origin of isolation.

CONCLUSION

It is clear based on the results of my study that samples of *Staphylococcus aureus* taken from medical clinics are not significantly differentiated from either hospital or community samples based on sequence data. Therefore, I conclude that medical environments – specifically ACCs – do not serve to generate uniquely adapted lineages. Moreover, my study has reiterated recent findings that *S. aureus* may be relatively homogenous genetically across a wide expanse of anthropogenic environments, as both hospital and community isolates appear indistinguishable from both clinics and one another.

The results of my study have implications from both a healthcare perspective as well as potential to inform future investigations into pathogen evolution. Given the high similarity of STs across clinics, community-carriage, and hospital samples, management efforts of *S. aureus* should be aware of the indistinct separation of these isolates. Specifically, managing *S. aureus* as 'community- or healthcare-associated' is not reflective of the genetic structure demonstrated by this and other studies (Lamers 2011). By delineating isolates into these two groups, it is possible that epidemiology of this species will not take into account the ability for strains to easily disperse in and out of healthcare settings, as strongly evidenced by MLST data. In short, the results of my study suggest that the monitoring and management of *S. aureus* should be more closely married across anthropogenic environments.

Additionally, my study will inform future investigations into *S. aureus* evolution. By addressing the potential hypothesis of lineage differentiation across anthropogenic environments, my study lays the foundation for future work in which alternative explanations for *S. aureus*

variation might arise. In particular, assessing the distribution of virulence factors, especially highly mobile elements such as resistance islands and prophages. Investigating these elements independently of *S. aureus* lineages may reveal patterns of inheritance that better explain the epidemiology of this species. Furthermore, the results of my study provide emphasis to the idea that pathology of this species may be driven by host factors. If lineages of *S. aureus* that are associated with hospital disease are indistinguishable from strains taken from healthy individuals in communities and clinics, as shown in my study, then it is plausible that pathogenesis is the result of maladaptive host traits. Further research into this area is necessary.

Finally, my study indicates the importance of continuing to develop new genetic markers for *S. aureus*. The weighted phylogeny's outperformance of the unweighted tree indicates that variable markers such as *spa* may strongly drive phylogenetic inference over less variable markers such as MLST if correct weighting parameters are not applied. Given a relative paucity of available markers and a reliance on MLST in previous studies, further investigation should prioritize enhancing the suite of available loci.

APPENDIX A: TABLES

Table 1: Demographic information of participants (a) and positive donors (b) by race. Asterisks denote positive numbers that were significantly different than expected (p<=.05)

Racial categories	Females	Males	Total	Percentage
White	49	25	74	52.48
Asian Pacific Islanders	8	3	11	7.80
Black or African American	22	8	30	21.28
American Indian or Alaskan Native	0	0	0	0.00
More than one race	6	1	7	4.96
Unknown	16	3	19	13.48
Total of all subjects	101	40	141	100.00

b)

Racial categories	Females	Males	Total	Percentage
White	13	6	19	65.5
Asian Pacific Islanders	4*	0*	4	13.7
Black or African American	1*	2	3	10.3
American Indian or Alaskan Native	0	0	0	0.00
More than one race	0*	0	0	0
Unknown	0*	3*	3	10.3
Total of all subjects	18	11	29	100.00

Table 2: Sequence of Primers utilized in PCR

Gene	Primer	Sequence (5'-3')
Carbamate kinase (*arcC*)	arcCF	TTGATTCACCAGCGCGTATTGTC
	arcCR	AGGTATCTGCTTCAATCAGCG
Shikimate dehydrogenase (*aroE*)	aroEF	ATCGGAAATCCTATTTCACATTC
	aroER	GGTGTTGTATTAATAACGATATC
Glycerol kinase (*glpF*)	glpF	CTAGGAACTGCAATCTTAATCC
	glpR	TGGTAAAATCGCATGTCCAATTC
Guanylate kinase (*gmk*)	gmkF	ATCGTTTTATCGGGACCATC
	gmkR	TCATTAACTACAACGTAATCGTA
Phosphate acetyltransferase (*pta*)	ptaF	GTTAAAATCGTATTACCTGAAGG
	ptaR	GACCCTTTTGTTGAAAAGCTTAA
Triosphosphate isomerase (*tpi*)	tpiF	TCGTTCATTCTGAACGTCGTGAA
	tpiR	TTTGCACCTTCTAACAATTGTAC
Acetyl coenzyme A Acetyltransferase (*yqiL*)	yqiLF	CAGCATACAGGACACCTATTGGC
	yqiLR	CGTTGAGGAATCGATACTGGAAC

Table 3: Partition of gene fragments and codon positions, with associated evolutionary models. Models include Felsenstein 1981 (F81), Hasegawa, Kishino, and Yano 1985 (HKY) and Generalised time-reversible (GTR, Tavare 1986). Variable site distributions are equal unless otherwise noted as proportion invariate (I) and Gamma distributions (G).

Gene Fragment	CP1	CP2	CP3
arcC	HKY+I+G	F81+I	HKY+G
aroE	HKY+I+G	F81+I	HKY+I+G
glpF	HKY+I+G	F81+I	HKY+I+G
gmk	HKY+I+G	F81+I	HKY+I+G
pta	HKY+I+G	F81+I	HKY+G
tpi	HKY+I+G	F81+I	HKY+I+G
yqiL	HKY+I+G	F81+I	HKY+I+G
spa	GTR+I	HKY+I	HKY+I

Table 4: Bayesian Factor Hypotheses testing of various competing hypotheses against the unconstrained (most likely) phylogeny. Bayesian factors exceeding a value of 150 are considered decisively significant – those exceeding 12 are strongly significant. Higher values indicate less likely hypotheses.

Hypothesis	BF (Stepping Stone)	BF (Path Sampling)
Constrained Hospital	399	400
Constrained Community	862	864
Constrained Clinical	852	854
Constrained Profession	163	163
Weighted	128	115

APPENDIX B: FIGURES

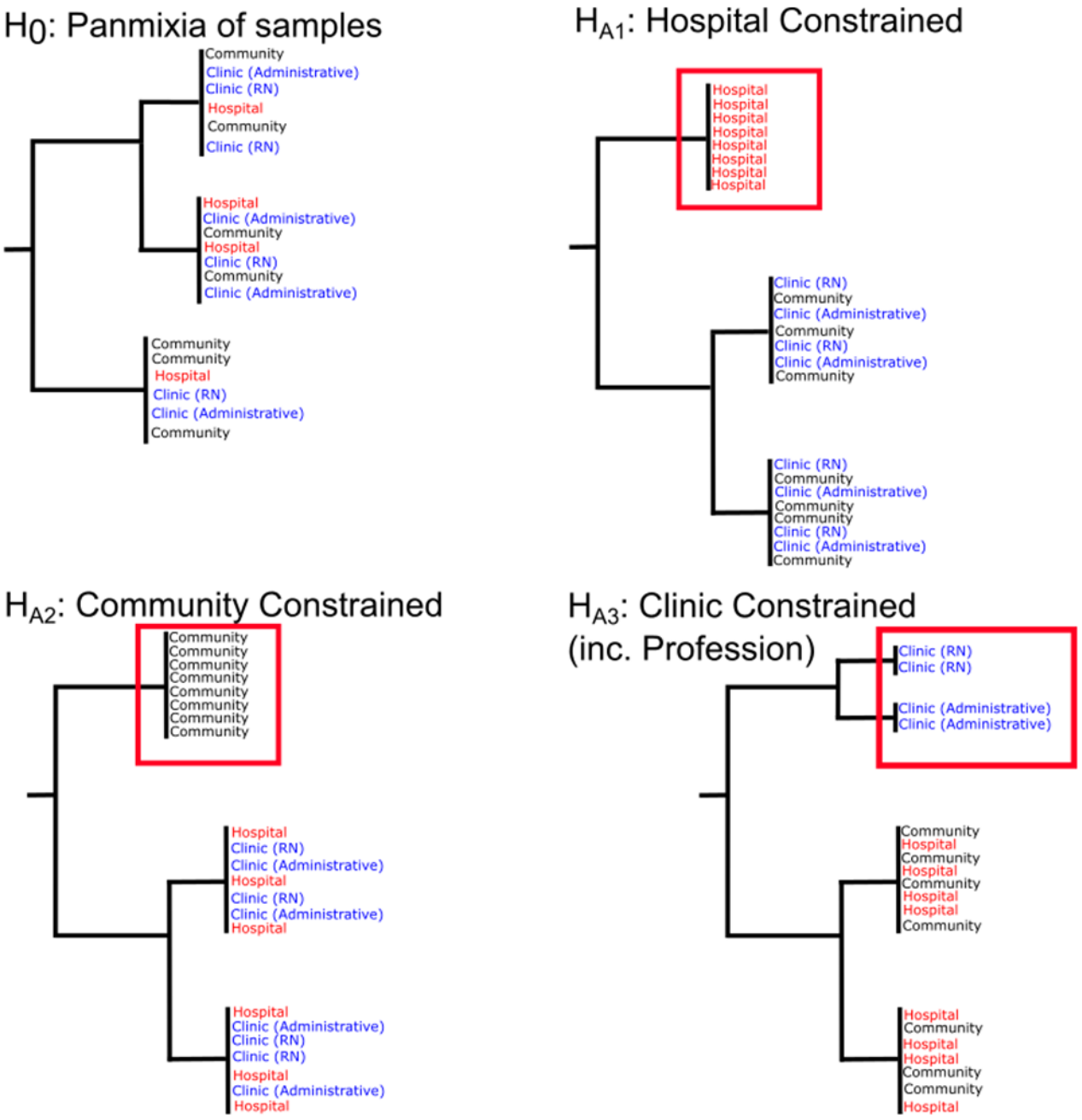

Figure 1: Schematic representation of all competing hypotheses as calculated utilizing Bayes Factor Testing. H_0 representing sample panmixia serves as the unconstrained reference tree that constrained alternative hypotheses are tested against.

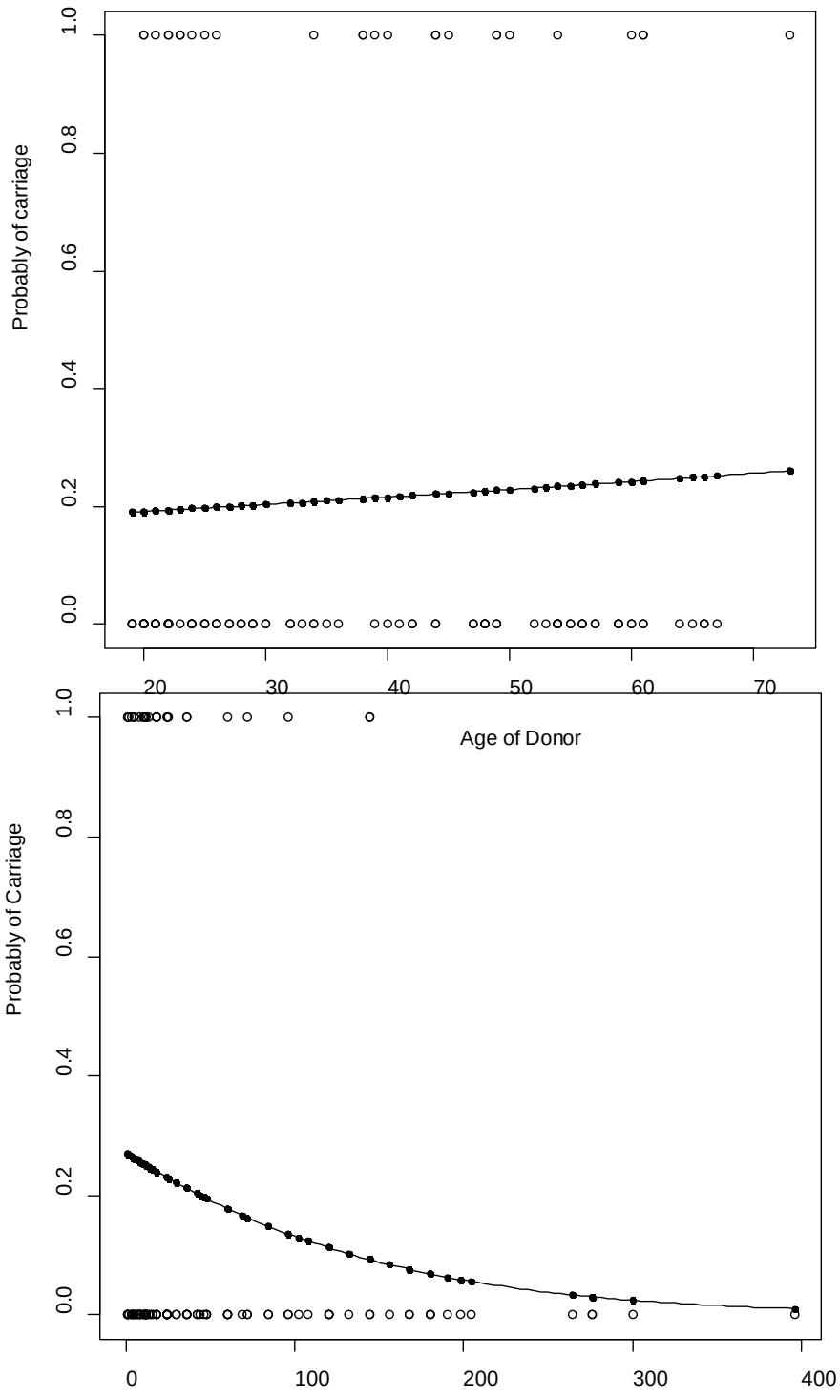

Figure 2: Logistic Regression of donor characteristics as predictors for carriage.

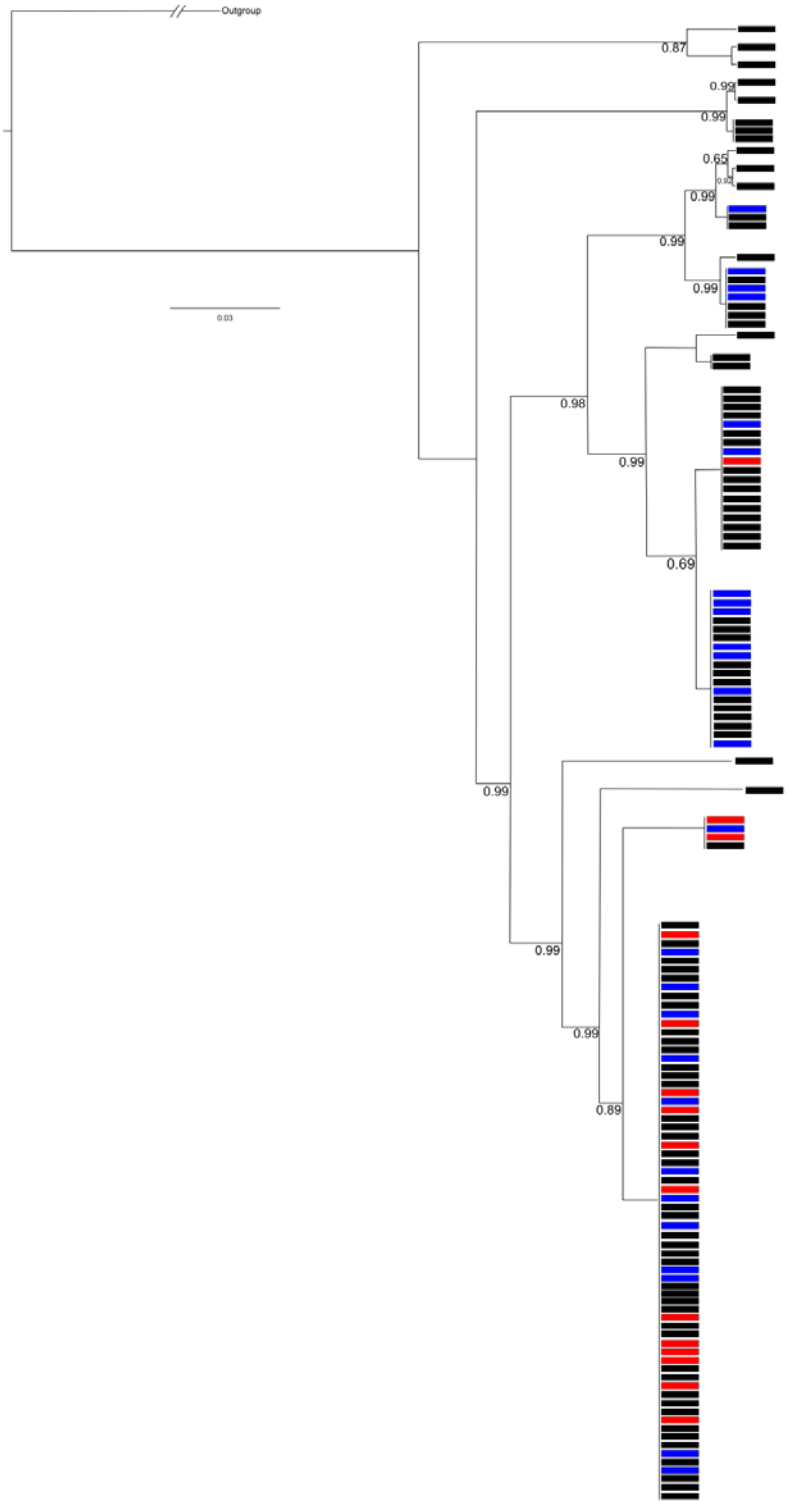

Figure 3: Most likely phylogenetic tree. Red taxa labels indicate NCBI hospital samples, blue taxon labels correspond to clinical samples, and black taxon to community samples. Solid circles on nodes indicate posterior probability values of 100.

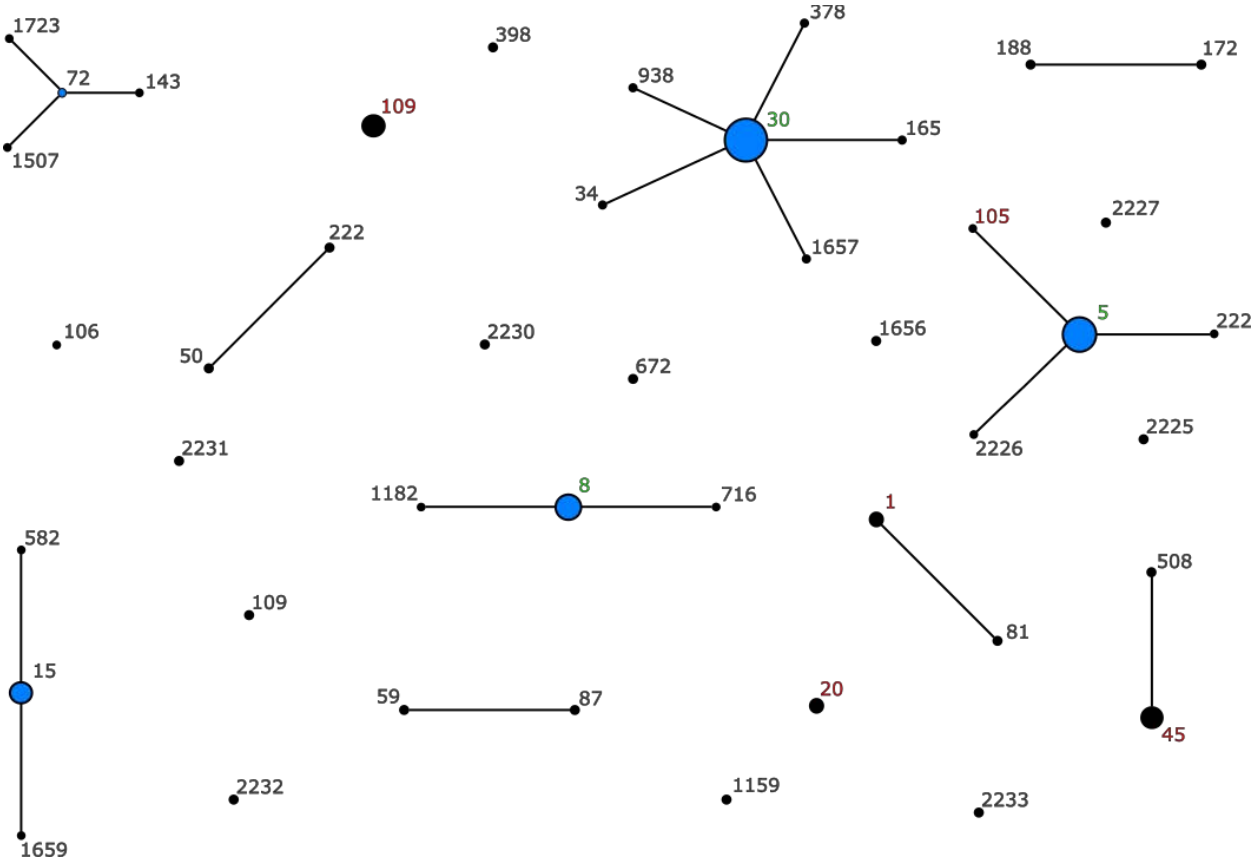

Figure 4: MLST EBURST of samples. Blue circles represent projected founder strains. Green labels represent strains shared between medical clinic and community cohorts. Black labels are community only, and red are clinical only. Size of circles are proportional to sampled number of isolates conforming to that strain.

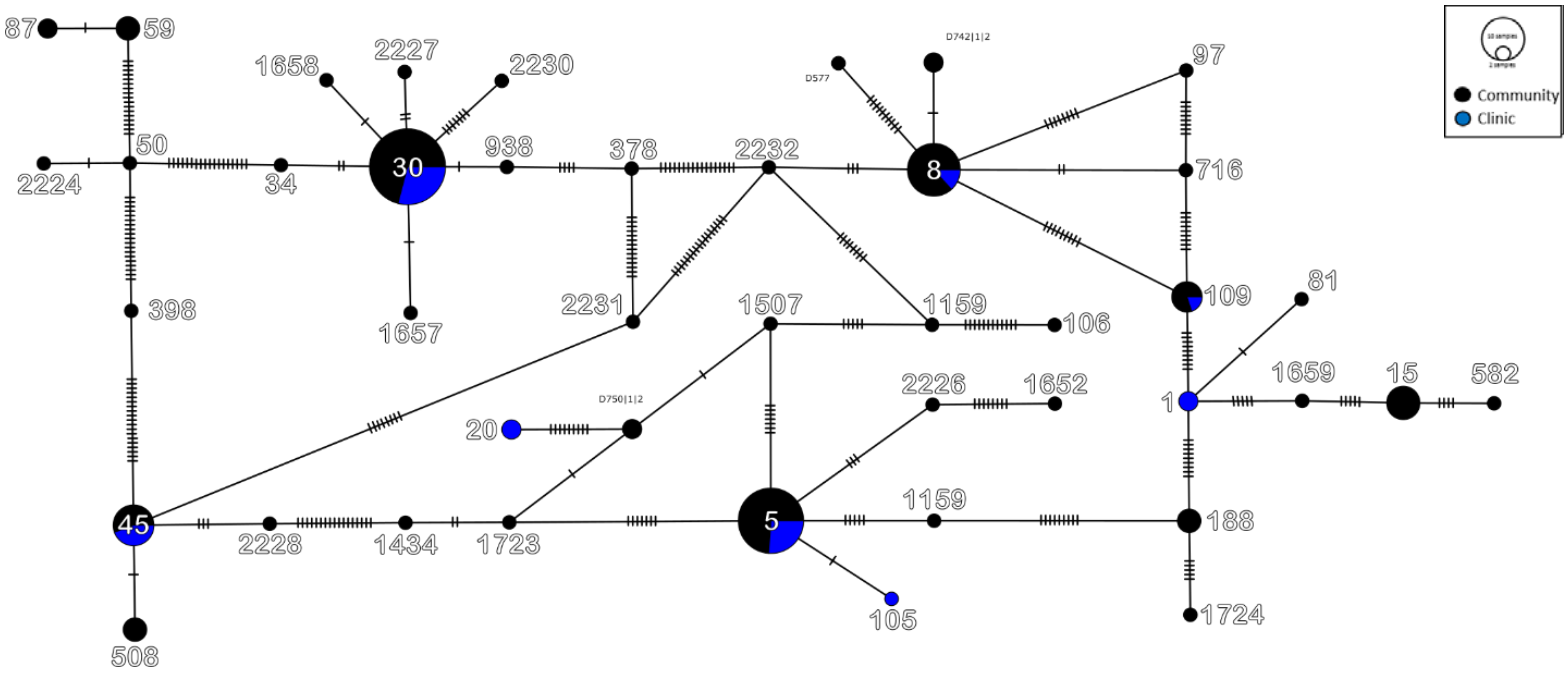

Figure 5: Minimal Spanning network of MLST data. Dashes represent inferred haplotypes not recovered in this sampling scheme. Numerical labels indicate the STs of isolates. Size of haplotype circle schematics are proportional to number of isolates recovered of that ST. Coloured wedges indicate proportion of ST owed to each sampled origin.

REFERENCES

Agnarsson, I., & May-Collado, L. J. (2008). The phylogeny of Cetartiodactyla: The importance of dense taxon sampling, missing data, and the remarkable promise of cytochrome b to provide reliable species-level phylogenies. Molecular Phylogenetics and Evolution, 48(3), 964–985. http://doi.org/10.1016/j.ympev.2008.05.046

Chambers, H. F. (2001). The changing epidemiology of Staphylococcus aureus? Emerging Infectious Diseases, 7(2), 178–182. http://doi.org/10.3201/eid0702.700178

Cheung, G. Y. C., Wang, R., Khan, B. a, Sturdevant, D. E., & Otto, M. (2011). Role of the accessory gene regulator agr in community-associated methicillin-resistant Staphylococcus aureus pathogenesis. Infection and Immunity, 79(5), 1927–35. http://doi.org/10.1128/IAI.00046-11

Chung, M., de Lencastre, H., Matthews, P., Tomasz, a, Adamsson, I., Aires de Sousa, M., … Villari, P. (2000). Molecular typing of methicillin-resistant Staphylococcus aureus by pulsed-field gel electrophoresis: comparison of results obtained in a multilaboratory effort using identical protocols and MRSA strains. Microbial Drug Resistance (Larchmont, N.Y.), 6(3), 189–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11144419

Clark, R. T., Hope, A., Lopez-Fraga, M., Schiller, N., & Lo, D. D. (2009). Bacterial particle endocytosis by epithelial cells is selective and enhanced by tumor necrosis factor receptor ligands. Clinical and Vaccine Immunology, 16(3), 397–407. http://doi.org/10.1128/CVI.00210-08

Cooper, J. E., & Feil, E. J. (2006). The phylogeny of Staphylococcus aureus - which genes make the best intra-species markers? Microbiology (Reading, England), 152(Pt 5), 1297–305. http://doi.org/10.1099/mic.0.28620-0

Cosgrove, S. E., Qi, Y., Kaye, K. S., Karchmer, A. W., & Carmeli, Y. (2013). Methicillin Resistance in Staphylococcus Aureus Bacteremia on Patient Outcomes : Mortality , Length of Stay , and Hospital Charges, 26(2), 517–523.

David, M. Z., & Daum, R. S. (2010). Community-Associated Methicillin-Resistant Staphylococcus aureus: Epidemiology and Clinical Consequences of an Emerging Epidemic. Clinical Microbiology Reviews, 23(3), 616–687. http://doi.org/10.1128/CMR.00081-09

Davis, K. a, Stewart, J. J., Crouch, H. K., Florez, C. E., & Hospenthal, D. R. (2004). Methicillin-resistant Staphylococcus aureus (MRSA) nares colonization at hospital admission and its effect on subsequent MRSA infection. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 39(6), 776–782. http://doi.org/10.1086/422997

Davis, S. L., Perri, M. B., Donabedian, S. M., Manierski, C., Singh, a., Vager, D., … Zervos, M. J. (2007). Epidemiology and Outcomes of Community-Associated Methicillin-Resistant Staphylococcus aureus Infection. Journal of Clinical Microbiology, 45(6), 1705–1711. http://doi.org/10.1128/JCM.02311-06

DeLeo, F. R., Otto, M., Kreiswirth, B. N., & Chambers, H. F. (2010). Community-associated meticillin-resistant Staphylococcus aureus. Lancet, 375(9725), 1557–68. http://doi.org/10.1016/S0140-6736(09)61999-1

Deurenberg, R. H., Vink, C., Kalenic, S., Friedrich, a W., Bruggeman, C. a, & Stobberingh, E. E. (2007). The molecular evolution of methicillin-resistant Staphylococcus aureus. Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases, 13(3), 222–35. http://doi.org/10.1111/j.1469-0691.2006.01573.x

Deurenberg, R. H., & Stobberingh, E. E. (2008). The evolution of Staphylococcus aureus. Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 8(6), 747–63. http://doi.org/10.1016/j.meegid.2008.07.007

Drummond, a. J., Suchard, M. a., Xie, D., & Rambaut, a. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution, 29(8), 1969–1973. http://doi.org/10.1093/molbev/mss075

Enright, M. C., & Spratt, B. G. (1999). Multilocus sequence typing. Trends in Microbiology, 7(12), 482–487. http://doi.org/10.1016/S0966-842X(99)01609-1

Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J., & Spratt, B. G. (2000). Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. Journal of Clinical Microbiology, 38(3), 1008–15.

Enright, M. C., Robinson, D. A., Randle, G., Feil, E. J., Grundmann, H., & Spratt, B. G. (2002). The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). Proceedings of the National Academy of Sciences of the United States of America, 99(11), 7687–92. http://doi.org/10.1073/pnas.122108599

Feil, E. J., Holmes, E. C., Bessen, D. E., Chan, M. S., Day, N. P., Enright, M. C., … Spratt, B. G. (2001). Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. Proceedings of the National Academy of Sciences of the United States of America, 98(1), 182–187. http://doi.org/10.1073/pnas.98.1.182

Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., & Spratt, B. G. (2004). eBURST: Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data. Journal of Bacteriology, 186(5), 1518–1530. http://doi.org/10.1128/JB.186.5.1518-1530.2004

Gillet, Y., Vanhems, P., Lina, G., Bes, M., Vandenesch, F., Floret, D., & Etienne, J. (2007). Factors predicting mortality in necrotizing community-acquired pneumonia caused by Staphylococcus aureus containing Panton-Valentine leukocidin. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 45(3), 315–321. http://doi.org/10.1086/519263

Grundmann, H., Aires-de-Sousa, M., Boyce, J., & Tiemersma, E. (2006). Emergence and resurgence of meticillin-resistant Staphylococcus aureus as a public-health threat. Lancet, 368(9538), 874–885. http://doi.org/10.1016/S0140-6736(06)68853-3

Harmsen, D., Claus, H., Witte, W., Claus, H., Turnwald, D., & Vogel, U. (2003). Typing of Methicillin-Resistant Staphylococcus aureus in a University Hospital Setting by Using Novel Software for spa Repeat Determination and Database Management, 41(12), 5442–5448. http://doi.org/10.1128/JCM.41.12.5442

Havird, J. C., & Miyamoto, M. M. (2010). The importance of taxon sampling in genomic studies: An example from the cyclooxygenases of teleost fishes. Molecular Phylogenetics and Evolution, 56(1), 451–455. http://doi.org/10.1016/j.ympev.2010.04.003

Holden, M. T. G., Feil, E. J., Lindsay, J. a, Peacock, S. J., Day, N. P. J., Enright, M. C., … Parkhill, J. (2004). Complete genomes of two clinical Staphylococcus aureus strains: evidence for the rapid evolution of virulence and drug resistance. Proceedings of the National Academy of Sciences of the United States of America, 101(26), 9786–9791. http://doi.org/10.1073/pnas.0402521101

Ito, T., Takeuchi, F., Okuma, K., Yuzawa, H., & Hiramatsu, K. (2004). Novel Type V Staphylococcal Cassette Chromosome. Society, 48(7), 2637–2651. http://doi.org/10.1128/AAC.48.7.2637

Karauzum, H., Ferry, T., de Bentzmann, S., Lina, G., Bes, M., Vandenesch, F., … Landmann, R. (2008). Comparison of adhesion and virulence of two predominant hospital-acquired methicillin-resistant Staphylococcus aureus clones and clonal methicillin-susceptible S. aureus isolates. Infection and Immunity, 76(11), 5133–8. http://doi.org/10.1128/IAI.01697-07

Katayama, Y., Ito, T., & Hiramatsu, K. (2000). A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in Staphylococcus aureus. Antimicrobial Agents and Chemotherapy, 44(6), 1549–1555. http://doi.org/10.1128/AAC.44.6.1549-1555.2000 Kaufmann, M. E. (1998), Pulsed-Field Gel Electrophoresis, Mol. Bacteriol. Protoc. Clin. Appl., 15, 33–50, doi:10.1385/0-89603-498-4:33.

Klevens, R. M., Morrison, M. a, Nadle, J., Petit, S., Gershman, K., Ray, S., … Fridkin, S. K. (2007). Invasive methicillin-resistant Staphylococcus aureus infections in the United States. JAMA : The Journal of the American Medical Association, 298(15), 1763–1771. http://doi.org/10.1001/jama.298.15.1763

Kluytmans, J., van Belkum, a, & Verbrugh, H. (1997). Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. Clinical Microbiology Reviews, 10(3), 505–20. http://doi.org/. doi:10.1016/S0006-3207(03)00146-0 Lamers, R. P., Stinnett, J. W., Muthukrishnan, G., Parkinson, C. L., & Cole, A. M. (2011). Evolutionary analyses of Staphylococcus aureus identify genetic relationships between nasal carriage and clinical isolates. PloS One, 6(1), e16426. http://doi.org/10.1371/journal.pone.0016426

Lamers, R. P., Muthukrishnan, G., Castoe, T. a, Tafur, S., Cole, A. M., & Parkinson, C. L. (2012). Phylogenetic relationships among Staphylococcus species and refinement of cluster groups based on multilocus data. BMC Evolutionary Biology, 12(1), 171. http://doi.org/10.1186/1471-2148-12-171

Lanfear, R., Calcott, B., Ho, S. Y. W., & Guindon, S. (2012). PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution, 29(6), 1695–1701. http://doi.org/10.1093/molbev/mss020

Lindsay, J. a. (2010). Genomic variation and evolution of Staphylococcus aureus. International Journal of Medical Microbiology, 300(2-3), 98–103. http://doi.org/10.1016/j.ijmm.2009.08.013

Lodise, T. P., McKinnon, P. S., Swiderski, L., & Rybak, M. J. (2003). Outcomes analysis of delayed antibiotic treatment for hospital-acquired Staphylococcus aureus bacteremia. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 36(11), 1418–1423. http://doi.org/10.1086/375057

Lorenz, E., Mira, J. P., Cornish, K. L., Arbour, N. C., & Schwartz, D. a. (2000). A novel polymorphism in the toll-like receptor 2 gene and its potential association with staphylococcal infection. Infection and Immunity, 68(11), 6398–6401. http://doi.org/10.1128/IAI.68.11.6398-6401.2000

Ma, X. X., Ito, T., Tiensasitorn, C., Jamklang, M., Chongtrakool, P., Boyle-Vavra, S., … Hiramatsu, K. (2002). Novel type of staphylococcal cassette chromosome mec identified in community-acquired methicillin-resistant Staphylococcus aureus strains. Antimicrobial Agents and Chemotherapy, 46(4), 1147–52. http://doi.org/10.1128/AAC.46.4.1147

Maiden, M. C., Bygraves, J. a, Feil, E., Morelli, G., Russell, J. E., Urwin, R., … Spratt, B. G. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America, 95(6), 3140–5.

McDougal, L. K., Steward, C. D., Killgore, G. E., Chaitram, J. M., McAllister, S. K., & Tenover, F. C. (2003). Pulsed-field gel electrophoresis typing of oxacillin-resistant Staphylococcus aureus isolates from the United States: establishing a national database. Journal of Clinical Microbiology, 41(11), 5113–20. http://doi.org/10.1128/JCM.41.11.5113

Montgomery, C. P., Boyle-Vavra, S., & Daum, R. S. (2010). Importance of the global regulators Agr and SaeRS in the pathogenesis of CA-MRSA USA300 infection. PloS One, 5(12), e15177. http://doi.org/10.1371/journal.pone.0015177

Murchan, S., Kaufmann, M. E., Deplano, A., de Ryck, R., Struelens, M., Zinn, C. E., … Cookson, B. (2003). Harmonization of pulsed-field gel electrophoresis protocols for epidemiological typing of strains of methicillin-resistant Staphylococcus aureus: a single approach developed by consensus in 10 European laboratories and its application for tracing the spre. Journal of Clinical Microbiology, 41(4), 1574–85. http://doi.org/10.1128/JCM.41.4.1574

Muthukrishnan, G., Lamers, R. P., Ellis, A., Paramanandam, V., Persaud, A. B., Tafur, S., … Cole, A. M. (2013). Longitudinal genetic analyses of Staphylococcus aureus nasal carriage dynamics in a diverse population. BMC Infectious Diseases, 13(1), 221. http://doi.org/10.1186/1471-2334-13-221

Ní Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Höök, M., Foster, T. J., & Ho, M. (1998). Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of Staphylococcus aureus. Molecular Microbiology, 30(2), 245–257. http://doi.org/10.1046/j.1365-2958.1998.01050.x

Noskin, G. a, Rubin, R. J., Schentag, J. J., Kluytmans, J., Hedblom, E. C., Jacobson, C., … Bharmal, M. (2007). National trends in Staphylococcus aureus infection rates: impact on economic burden and mortality over a 6-year period (1998-2003). Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 45(9), 1132–40. http://doi.org/10.1086/522186

Nouwen, J. L., Ott, A., Kluytmans-Vandenbergh, M. F. Q., Boelens, H. a M., Hofman, A., van Belkum, A., & Verbrugh, H. a. (2004). Predicting the Staphylococcus aureus nasal carrier state: derivation and validation of a “culture rule”. Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 39(6), 806–811. http://doi.org/10.1086/423376

Ogston A. (1882). Micrococcus poisoning. J Anat;17:24-58.

Patel, N. N., & Patel, D. N. (2010). Staphylococcal Scalded Skin Syndrome. The American Journal of Medicine, 123(6), 505–507. http://doi.org/10.1016/j.amjmed.2009.09.041
Posada, D., & Buckley, T. (2004). Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. Systematic Biology, 53(5), 793–808. http://doi.org/10.1080/10635150490522304
Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. Biometrika, 83, 251–266.
Rammelkamp, C.H., & Maxon, T. (1942). Resistance of Staphylococcus aureus to the Action of Penicillin. Experimental Biology and Medicine, 51(3), pp.386-389.
Rieg, G., Mehdi, S., Perlroth, J., Bayer, A. S., Tang, A. W., Phung, T. O., & Spellberg, B. (2005). Necrotizing Fasciitis Caused by Community-Associated Methicillin-Resistant. N Engl J Med, 352(14), 1445–1453. http://doi.org/10.1056/NEJMoa042683
Rohde, H., Burandt, E. C., Siemssen, N., Frommelt, L., Burdelski, C., Wurster, S., … Mack, D. (2007). Polysaccharide intercellular adhesin or protein factors in biofilm accumulation of Staphylococcus epidermidis and Staphylococcus aureus isolated from prosthetic hip and knee joint infections. Biomaterials, 28(9), 1711–1720. http://doi.org/10.1016/j.biomaterials.2006.11.046
Rotun, S. S., McMath, V., Schoonmaker, D. J., Maupin, P. S., Tenover, F. C., Hill, B. C., & Ackman, D. M. (1999). Staphylococcus aureus with reduced susceptibility to vancomycin isolated from a patient with fatal bacteremia. Emerging Infectious Diseases, 5(1), 147–149. http://doi.org/10.3201/eid0501.990118
Schappert, S. M. (1998). Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments: United States, 1996. Vital and Health Statistics. Series 13, Data from the National Health Survey, (157), 1–37. http://doi.org/http://europepmc.org/abstract/MED/16471269
Shuter, J., & Hatcher, V. B. (1996). Staphylococcus aureus binding to human nasal These include : Staphylococcus aureus Binding to Human Nasal Mucin, 64(1), 310–318.
Slatkin, M., & Maddison, W. P. (1990). Detecting isolation by distance using phylogenies of genes. Genetics, 126(1), 249–260.
Starfield, B., Weiner, J., Mumford, L., & Steinwachs, D. (1991). Ambulatory care groups: a categorization of diagnoses for research and management. Health Services Research, 26(1), 53–74.
Tabaqchali, S. (1997). Vancomycin-resistant Staphylococcus aureus : apocalypse now ? European perspectives on general medicine. The Lancet, 350, 1644–1645. http://doi.org/10.1016/S0140-6736(05)64269-9
Tancrede, C. (1992), Role of human microflora in health and disease, Eur. J. Clin. Microbiol. Infect. Dis., 11(11), 1012–1015, doi:10.1007/BF01967791.
Tenover, F. C., Weigel, L. M., Appelbaum, P. C., Mcdougal, L. K., Chaitram, J., Mcallister, S., … Patel, J. B. (2004). Vancomycin-Resistant Staphylococcus aureus Isolate from a Patient in Pennsylvania, 48(1), 275–280. http://doi.org/10.1128/AAC.48.1.275
Tsiodras, S., Gold, H. S., Sakoulas, G., Eliopoulos, G. M., Wennersten, C., Venkataraman, L., … Ferraro, M. J. (2001). Linezolid resistance in a clinical isolate of Staphylococcus aureus. Lancet, 358(9277), 207–8. http://doi.org/10.1016/S0140-6736(01)05410-1
van Belkum, A., Verkaik, N. J., de Vogel, C. P., Boelens, H. a, Verveer, J., Nouwen, J. L., … Wertheim, H. F. L. (2009). Reclassification of Staphylococcus aureus nasal carriage types. The Journal of Infectious Diseases, 199(12), 1820–6. http://doi.org/10.1086/599119
Vandenesch, F., Naimi, T., Enright, M. C., Lina, G., Nimmo, G. R., Heffernan, H., … Etienne, J. (2003). Community-acquired methicillin-resistant staphylococcus aureus carrying panton-valentine leukocidin genes: Worldwide emergence. Emerging Infectious Diseases, 9(8), 978–984. http://doi.org/10.3201/eid0908.030089
Voyich, J. M., Otto, M., Mathema, B., Braughton, K. R., Whitney, A. R., Welty, D., … DeLeo, F. R. (2006). Is Panton-Valentine leukocidin the major virulence determinant in community-associated methicillin-resistant Staphylococcus aureus disease? The Journal of Infectious Diseases, 194(12), 1761–1770. http://doi.org/10.1097/01.inf.0000256370.53802.a6
Wertheim, H. F. L., Melles, D. C., Vos, M. C., Leeuwen, W. Van, Belkum, A. Van, Verbrugh, H. A., & Nouwen, J. L. (2005). Review The role of nasal carriage in Staphylococcus aureus infections, 5(December), 751–762.

Williams, R. E. (1963). Healthy carriage of Staphylococcus aureus: its prevalence and importance. Bacteriological Reviews, 27(96), 56–71.

Young, B. C., Golubchik, T., Batty, E. M., Fung, R., Larner-Svensson, H., Votintseva, A. a, … Wilson, D. J. (2012). Evolutionary dynamics of Staphylococcus aureus during progression from carriage to disease. Proceedings of the National Academy of Sciences of the United States of America, 109(12), 4550–5. http://doi.org/10.1073/pnas.1113219109

Zhang, K., McClure, J. A., & Conly, J. M. (2012). Enhanced multiplex PCR assay for typing of staphylococcal cassette chromosome mec types I to V in methicillin-resistant Staphylococcus aureus. Molecular and Cellular Probes, 26(10), 218–221. http://doi.org/10.1016/j.mcp.2012.04.002